DOUTORAMENTO
BIOLOGIA MOLECULAR E CELULAR

Decoding Heparan Sulfate unique structural epitopes involved in cancer cell signaling

Catarina Marques

D

2024

Decoding Heparan Sulfate unique structural epitopes involved in cancer cell signaling

Catarina Marques

TESE DE DOUTORAMENTO APRESENTADA AO INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR DA UNIVERSIDADE DO PORTO EM BIOLOGIA MOLECULAR E CELULAR E À ESCOLA DOUTORAL "*CSV- CHIMIE ET SCIENCES DU VIVANT*" DA UNIVERSIDADE DE GRENOBLE-ALPES NA ESPECIALIDADE "*CHIMIE BIOLOGIE*"

CATARINA MARQUES

Decoding Heparan Sulfate unique structural epitopes involved in cancer cell signaling

Tese de Candidatura ao grau de Doutor em Biologia Molecular e Celular submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientadora – Professora Doutora Ana Maria Rodrigues Leite de Magalhães
Categoria – Professora Associada
Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

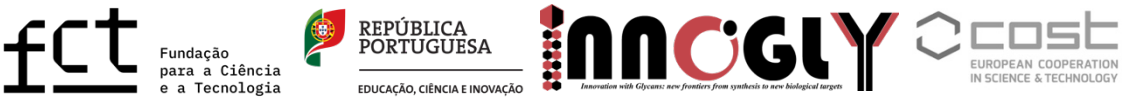
Coorientador –Doutor Romain R. Vivès
Categoria – French National Centre for Scientific Research (CNRS) Research Director
Afiliação – Institut de Biologie Structurale

Coorientador – Professor Doutor Celso Albuquerque Reis
Categoria – Professor Associado Convidado
Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Financiamento / Funding

O trabalho aqui apresentado foi financiado pela bolsa de doutoramento com a referência *2020.06412.BD (https://doi.org/10.54499/2020.06412.BD)* e pelo projeto PTDC/MED-ONC/28489/2017 atribuídos pela Fundação Portuguesa para a Ciência e a Tecnologia (FCT) do Ministério da Ciência, Tecnologia e Ensino Superior, e por duas bolsas destinadas a missões científica a curto prazo atribuídas pela INNOGLY COST Action (Action number: CA18103).

The work presented within this thesis was financially supported by the PhD fellowship with the reference *2020.06412.BD (https://doi.org/10.54499/2020.06412.BD)* and the research project PTDC/MED-ONC/28489/2017 granted by the Portuguese Foundation for Science and Technology (FCT) of the Portuguese Ministry of Science, Technology and Higher Education, and by two Short-Term Scientific Mission Grants awarded by the INNOGLY COST Action (Action number: CA18103).



-Vivemos entre os homens, ajudemos os homens.

-E que faz o senhor para isso?

-Conserto-lhes os sapatos, já que nada mais posso fazer agora.

Claraboia, José Saramago

List of Publications

This thesis is based on the following publications:

Marques C, Poças J, Gomes C, Faria-Ramos I, Reis CA, Vivès RR, Magalhães A. Glycosyltransferases EXTL2 and EXTL3 cellular balance dictates heparan sulfate biosynthesis and shapes gastric cancer cell motility and invasion. *J Biol Chem.* (2022); doi: 10.1016/j.jbc.2022.102546.

Marques C, Reis CA, Vivès RR, Magalhães A. Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression. *Front Oncol.* (2021); doi: 10.3389/fonc.2021.778752.

In preparation:

Marques C, Ribeiro AR, Gomes C, Abrantes R, Gout E, Cavadas B, Reis CA, Vivès RR, Magalhães A. Disclosing the regulatory role of Heparan Sulfate 6-*O*-sulfation patterns in gastric cancer. (*In preparation*).

Marques C, Ribeiro AR, Pinto F, Cavadas B, Barros R, Carneiro F, Reis CA, Vivès RR, Magalhães A. Uncovering the Gastric Cancer Glycosaminoglycan Landscape. (*In preparation*).

Other publications:

Poças J, **Marques C**, Gomes C, Otake A H, Pinto F, Ferreira M, Silva T, Magalhães A, et al. Syndecan-4 is a maestro of gastric cancer cell invasion and communication that underscores poor survival. *Proc Natl Acad Sci USA* (2023); doi: 10.1073/pnas.2214853120.

Faria-Ramos I, Poças J, **Marques C**, Santos-Antunes J, Macedo G, Carneiro F, Reis CA, Magalhães A. Heparan sulfate glycosaminoglycans: (un)expected allies in cancer clinical management. *Biomolecules*. (2021); doi: 10.3390/biom11020136.

Agradecimentos / Acknowledgements

Gostaria de começar por agradecer à minha orientadora, a Dra. Ana Magalhães, que me acompanhou nesta bonita jornada com uma dedicação e generosidade incomparáveis. À Ana, agradeço pelas valiosas discussões científicas que partilhámos, pela confiança que depositou em mim e pela liberdade que me proporcionou ao longo deste projeto. Através do seu próprio exemplo, ensinou-me trabalho árduo e perseverança, motivando-me a dar o meu melhor, sempre com otimismo e um entusiasmo contagiantes pela ciência. No fundo, agradeço pelo tanto que me ensinou, pelos novos "mundos" que me deu e pela paciência e tempo que me dedicou, sem os quais não teria chegado aqui.

I am also incredibly grateful to Dr. Romain Vivès for accepting to be my co-supervisor and for all his support. Despite being far away, Romain diligently took on the role of an attentive co-supervisor. I would also like to thank him for welcoming me to the "Structure et Activité des Glycosaminoglycanes" lab and for his scientific guidance during my stay at Grenoble, particularly marked by his good humor and enthusiasm for science.

In addition, I would also like to thank all the members of the "Structure et Activité des Glycosaminoglycanes" group for also kindly welcoming me and supporting me during this period. The people I found in Grenoble really taught me the true meaning of a "home away from home", their kindness and cheerful spirit lightened every darker moment, and I am truly greatfull for their friendship.

Ao Dr. Celso Reis, agradeço também por ter aceitado ser meu co-orientador, pela oportunidade de realizar o meu projeto de doutoramento no grupo de investigação "Glycobiology in Cancer", bem como por todo o apoio prestado ao longo destes anos. O seu empenho e dedicação pela ciência, e em particular pelo "Glycobiology in Cancer" são admiráveis, e contribuíram sem sombra de dúvida para o meu crescimento pessoal e científico.

Um sentido agradecimento a todos os membros do grupo "Glycobiology em Cancer" que tive a sorte de conhecer, não só pela partilha científica, espírito altruísta, que foram sem dúvida cruciais neste percurso, mas também pela amizade, compreensão e companheirismo. À Dra. Catarina Gomes agradeço pela enorme disponibilidade e apoio científico que me deu ao longo destes anos, que foram cruciais para o desenvolvimento de vários projetos e para o meu crescimento científico. À Juliana Poças e Isabel Faria-Ramos, um especial obrigado pelo apoio imenso que me deram no início do meu percurso e pela infindável paciência que tiveram comigo quando a ciência e o laboratório me eram ainda pouco familiares. À Raquel Ribeiro, agradeço pelo companheirismo e amizade que partilhamos nos mais recentes anos, e pelo apoio e otimismo constantes. A Raquel trouxe

consigo uma inteligência e perseverança fora do comum, e teve a generosidade de partilhar o seu percurso comigo, fazendo também ela parte do meu.

Aos meus amigos mais antigos, Catarina, Joana, Mariana e Tiago, e às amigas de sempre que a faculdade me deu, Anastasiya, Cláudia, Rafaela e Telma, agradeço pela alegria e o apoio constantes que são na minha vida.

Obrigada também à minha família, por vezes barulhenta, teimosa e opinativa, mas sempre presente, com o maior abraço do mundo à distância de uma chamada.

À minha mãe agradeço com muita admiração pela coragem com que me educou e trilhou este caminho de mãos dadas comigo. Obrigada mãe pelo amor, liberdade e amparo incondicionais, e por acreditares sempre em mim, vendo para lá das minhas dúvidas e medos.

Por fim, um enorme obrigado ao Leo, por tudo o que é e pelo tanto que me faz feliz, e à entidade ou acaso do destino que nos colocou no caminho um do outro.

Abstract

Gastric cancer persists as an aggressive and highly mortal pervasive pathology, mostly due to its asymptomatic nature, limited screening strategies for early detection and scarce treatment options for patients in advanced stages. Aberrant glycosylation is a hallmark of cancer, reported in multiple types of malignancies, therefore holding great potential for the discovery of novel screening tools and therapeutic targets.

Heparan sulfate (HS) proteoglycans (HSPGs) are fundamental glycoconjugates of the cell glycocalyx and extracellular matrix (ECM), that play key roles in physiology and pathology. HSPGs mainly function as cell surface scaffolds for receptor-ligand interactions and as ECM storage units, by binding and regulating the gradient of extracellular molecules, overall fine-tuning multiple cellular signaling networks and cell-cell/ECM crosstalk. HS structural features, including epimerization and sulfation, impact HSPGs-protein binding affinity, and consequently their biological roles. Several HSPGs and enzymes that catalyze fundamental steps in HS biosynthesis are abnormally expressed in a vast diversity of cancer pathologies, including gastric cancer. These molecular changes influence tumor cell communication and motility, and modulate the tumor microenvironment. Nonetheless, the specific tumorigenic steps guided by HS chains in gastric cancer remain to be fully elucidated. In this thesis, our principal objective was to understand the regulatory mechanisms underlying HS biosynthesis and cellular HSPGs remodeling in gastric cancer, and to disclose the oncogenic functions encoded by HS chains.

To explore the role of HS in cancer cell glycoproteome, signaling networks and motility, we established glycoengineered gastric cancer cell models either lacking important HS glycosyltransferases, or harboring altered expression of HS 6-*O*-sulfation regulatory enzymes.

Transcriptomic and disaccharide analyses showed that Exostosin-Like 3 (EXTL3) is key in initiating HS biosynthesis in gastric cancer cells, and that it holds an important role in modulating cellular HS/CS ratio. Conversely, we demonstrated that EXTL2 is an inhibitory regulator of HS synthesis, and that its abrogation promoted the biosynthesis of HS, with decreased 6-*O*-sulfation motifs, and upregulated Syndecan-4 (SDC4), a major cell-surface HSPG associated with gastric carcinogenesis. Functional assays revealed that this aberrant glycosylation profile increased cell motility and invasion and impaired the activation of cell surface receptors tyrosine kinase.

Moreover, we revealed that ablation of intracellular HS 6-*O*-sulfation, through the KO of the sulfotransferase HS6ST1, induced HS biosynthesis along with SDC4 expression. In opposition, the overexpression of the extracellular 6-*O*-endosulfatase HSULF2, which also impacted HS 6-*O*-sulfation, did not elicit the same HS/HSPG changes. We further revealed

that HSULF2 is highly expressed in gastric cancer and associates with patient's worse prognosis.

Finally, we analyzed the transcriptomic profile of several glycosaminoglycan (GAG)-related genes using six gastric cancer cell lines and samples from a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD). These data outlined the cell-type specificity and importance of GAG biosynthesis in gastric cancer, and further supported that deregulation of the HS/CS metabolic system significatively impacts cancer progression and patient prognosis.

Overall, our work has provided novel insights on the pro-tumorigenic roles of HS in gastric cancer, and has further proposed a regulatory network combining HS biosynthesis and structural motifs with the expression and glycosylation of SDC4, an important tumor promoter in the gastric context. These data highlight the clinical potential of HS and HS biosynthesis enzymes as targets in cancer therapy, and sets the foundation for the discovery of new biomarkers for early disease detection and patient stratification.

Keywords: Proteoglycans; Heparan Sulfate; Glycosaminoglycans; Gastric Cancer.

Resumo

O cancro gástrico persiste como uma patologia agressiva e extremamente letal, devido principalmente à sua natureza assintomática, às limitadas estratégias de rastreio que visam a deteção precoce da doença e às escassas opções de tratamento para os doentes em fase avançada. A glicosilação aberrante é uma caraterística distinta associada a cancro, descrita em vários tipos de neoplasias malignas, pelo que apresenta um grande potencial para a descoberta de novas ferramentas de rastreio e alvos terapêuticos.

Os proteoglicanos de sulfato de heparano (*HSPGs*) são glicoconjugados fundamentais do glicocálice das células e da matriz extracelular (*ECM*) que desempenham papéis fundamentais em condições fisiológicas e patológicas. Os *HSPGs* modulam interações recetor-ligando na superfície celular, e funcionam como unidades de armazenamento da *ECM*, ligando-se a moléculas e regulando o seu gradiente de concentração extracelular, controlando assim várias redes de sinalização celular e a interação célula-célula/*ECM*. As caraterísticas estruturais do *HS*, incluindo a epimerização e a sulfatação, têm impacto na afinidade de ligação dos *HSPGs* às proteínas e, consequentemente, nas suas funções biológicas. Vários *PGs* e enzimas que catalisam passos fundamentais na biossíntese de *HS* têm expressão alterada em variados tipos de cancro, incluindo o cancro gástrico. Estas alterações moleculares influenciam a comunicação e a motilidade das células tumorais, e modulam o microambiente tumoral. Porém, as vias específicas guiadas pelas cadeias de *HS* na formação e progressão de cancro gástrico ainda não são completamente conhecidas. Nesta tese, tivemos como principal objetivo compreender os mecanismos celulares reguladores subjacentes à biossíntese de *HS* e à remodelação de *HSPGs* em cancro gástrico, e revelar as funções pró-oncogénicas codificadas nas cadeias de *HS*.

Com o objetivo de explorar o papel do *HS* na formação do glicoproteoma das células tumorais, assim como nas vias de sinalização e na motilidade, recorremos à engenharia genética e estabelecemos modelos celulares de cancro gástrico que carecem de importantes glicosiltransferases de *HS* ou que apresentam expressão alterada de enzimas de regulação da 6-*O*-sulfatação de *HS*.

Análises de transcriptómica e de estrutura de dissacáridos revelaram que a enzima *Exostosin-Like 3* (EXTL3) é fundamental para iniciar a biossíntese de *HS* nas células de cancro gástrico e que desempenha um papel importante na modulação do rácio celular de *HS/CS*. Por outro lado, a EXTL2 é uma enzima reguladora inibitória da síntese de *HS*, e a sua eliminação promoveu a biossíntese de HS, com diminuição dos motivos de 6-*O*-sulfatação, e aumentou a expressão de *Syndecan 4* (SDC4), um importante *HSPG* na superfície das células, previamente associado à carcinogénese gástrica. Ensaios funcionais revelaram que este perfil alterado de glicosilação aumentou a motilidade e a

invasão celular e inibiu a ativação de recetores de tirosina quinase (RTKs) na superfície celular.

Adicionalmente, revelámos que a eliminação da 6-*O*-sulfatação intracelular nas cadeias de HS, através do *KO* da sulfotransferase HS6ST1, aumentou a biossíntese de *HS* juntamente com a expressão de *SDC4*. Em oposição, a sobreexpressão da 6-*O*-endossulfatase extracelular, HSULF2, que também tem impacto na 6-*O*-sulfatação de *HS*, não provocou as mesmas alterações em ternos de níveis de *HS/HSPG*. Revelámos ainda que a expressão da *HSULF2* está elevada em cancro gástrico e associada ao pior prognóstico dos pacientes.

Por último, analisámos o perfil de transcriptómica de vários genes relacionados com os glicosaminoglicanos (GAG) utilizando seis linhas celulares de cancro gástrico e amostras de tecidos de pacientes de cancro gástrico do *The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD)*. Estes dados demonstraram a importância da biossíntese de GAGs em cancro gástrico, dependente do tipo celular, e mostraram que a desregulação do sistema metabólico *HS*/Sulfato de Condroitina (*CS*) tem um impacto significativo na progressão do cancro e no prognóstico dos pacientes.

Em síntese, o nosso trabalho gerou novo conhecimento sobre os papéis promotores de tumorigénese do *HS* em cancro gástrico e propôs a existência de uma via de regulação que envolve a biossíntese e motivos estruturais específicos de *HS* e a expressão e glicosilação do SDC4, uma molécula relevante no contexto gástrico. Estes dados realçam o potencial das enzimas de biossíntese de *HS* e do próprio *HS* como alvos na terapia do cancro e estabelecem as bases moleculares para a descoberta de novos biomarcadores de deteção precoce da doença e estratificação dos pacientes.

Palavras-chave: Proteoglicanos; Sulfato de Heparano; Glicosaminoglicanos; Cancro Gástrico**.**

Résumé

Le cancer gastrique reste une pathologie agressive et hautement mortelle, principalement en raison de sa nature asymptomatique, des stratégies de dépistage limitées pour la détection précoce et des rares options de traitement pour les patients à un stade avancé. La glycosylation aberrante est une caractéristique du cancer, reportée dans de nombreux types de tumeurs malignes, et présente donc un fort potentiel pour la découverte de nouveaux outils de dépistage et de nouvelles cibles thérapeutiques.

Les protéoglycanes à Héparane Sulfate (HSPG) sont des glycoconjugués importants du glycocalyx cellulaire et de la matrice extracellulaire (MEC), qui jouent un rôle clé dans de nombreux processus physiologiques et pathologiques. Les HSPG servent principalement d'échafaudages à la surface des cellules pour faciliter les interactions récepteur-ligand, ou de réservoirs de stockage dans la MEC, capables de capturer et de réguler la formation de gradients de molécules extracellulaires. Ces larges propriétés d'interaction contribuent à réguler avec précision de multiples réseaux de signalisation cellulaire et les interactions cellule-cellule/MEC. Les caractéristiques structurales des HS, notamment leur degré d'épimérisation et de sulfatation, ont un impact sur l'affinité de leurs interactions avec les protéines et, par conséquent, sur leurs fonctions biologiques. Ainsi, plusieurs HSPG et enzymes de biosynthèse des HS sont exprimées de manière anormale dans de nombreux cancers, notamment le cancer gastrique. Ces changements moléculaires influencent la communication et la motilité des cellules tumorales et modulent le microenvironnement tumoral. Néanmoins, les étapes tumorigènes spécifiques associées aux HS dans le cancer gastrique restent à élucider. Dans cette thèse, notre objectif principal était de comprendre les mécanismes de régulation qui sous-tendent la biosynthèse des HS et le remodelage des HSPGs cellulaires dans le cancer gastrique, et de révéler les fonctions oncogéniques controlées par les HS.

Pour explorer l'implication des HS dans le glycoprotéome des cellules cancéreuses, les réseaux de signalisation et la motilité, nous avons établi des modèles de cellules cancéreuses gastriques glyco-ingéniées, soit dépourvues de glycosyltransférases importantes pour la synthèse d'HS, soit présentant une expression altérée d'enzymes régulatrices de la 6-*O*-sulfatation des HS.

Des études transcriptomiques couplées à des analyses de composition disaccharidique des glycosaminoglycanes (GAG) ont montré que l'Exostosine-Like 3 (EXTL3) joue un rôle clé dans l'initiation de la biosynthèse des HS dans les cellules cancéreuses gastriques, et qu'elle joue un rôle important dans la modulation du rapport HS/Chondroitine Sulfate (CS) cellulaire. À l'inverse, EXTL2 est un régulateur inhibiteur de la synthèse des HS, et son abrogation favorise la biosynthèse des HS, avec une diminution des motifs de 6-*O*-sulfation, et une régulation à la hausse du Syndécan-4 (SDC4), un HSPG

majeur de la surface cellulaire associé à la carcinogenèse gastrique. Des essais fonctionnels ont révélé que ce profil de glycosylation aberrant augmentait la motilité et l'invasion des cellules et entravait l'activation des récepteurs tyrosine kinase de la surface cellulaire.

De plus, nous avons révélé que la repression de la 6-*O*-sulfatation des HS, par l'inhibition de la sulfotransférase HS6ST1, augmentait la synthèse d'HS ainsi que l'expression du SDC4. En revanche, la surexpression de la 6-*O*-endosulfatase extracellulaire HSULF2, qui a également un impact sur la 6-*O*-sulfatation des HS, n'a pas les mêmes effets sur les niveaux d'HS/HSPG exprimés. Nous avons également montré que HSULF2 est fortement exprimé dans le cancer gastrique et qu'il est associé à un pronostic plus défavorable pour les patients.

Enfin, nous avons analysé le profil transcriptomique de plusieurs gènes liés aux GAG en utilisant six lignées cellulaires de cancer gastrique et des échantillons d'une cohorte de patients atteints de cancer gastrique provenant de l'Atlas du génome du cancer de l'adénocarcinome de l'estomac (TCGA-STAD). Ces données soulignent la spécificité du type cellulaire et l'importance de la biosynthèse des GAG dans le cancer gastrique, et confirment qu'une dérégulation de l'expression et de la structure des HS/CS a un impact significatif sur la progression du cancer et le pronostic des patients.

Dans l'ensemble, nos travaux ont fourni de nouvelles informations sur les rôles pro-tumorigènes des HS dans le cancer gastrique et mettent en évidence l'existence d'un réseau de régulation impliquant la biosynthèse des HS, la formation de motifs structuraux et l'expression et la glycosylation du SDC4, un HPS aux propriétés pro-tumorales dans le contexte du cancer gastrique. Ces données soulignent le potentiel clinique des HS et des enzymes de biosynthèse des HS en tant que cibles thérapeutiques anticancéreuses et jettent les bases de la découverte de nouveaux biomarqueurs pour la détection précoce de la maladie et la stratification des patients.

Mots-clés: Protéoglycanes; Héparane Sulfate; Glycosaminoglycanes; Cancer de l'estomac.

Abbreviations

AJCC - American Joint Committee on Cancer

BMP - Bone Morphogenetic Proteins

c-Met - Hepatocyte growth factor receptor

C1GalT1 - Gal-transferase

C2GnT - core 2 β1-6 *N*-acetylglucosaminyltransferase

CagA - Cytotoxin-associated gene A

Cas9 – CRISPR-associated protein 9

CDH1 - E-cadherin

CEA - Carcinoembryonic antigen

CHPF - CS synthase 2

CHPF2 - CS glucuronyltransferase

CHST14 - Dermatan 4-O-sulfotransferase

CHST15 - GalNAc 4-O-sulfate 6-O-sulfotransferase

CHSY1 - CS synthase 1

CHSY3 - CS synthase 3

CIN - Chromosomal Instability

CRISPR-Cas9 - Clustered regularly interspaced short palindromic repeats/Cas9

CS - Chondroitin sulfate

CSGALNACT1 and 2 - CS *N*-acetylgalactosaminyltransferase 1 and 2

DFS – Disease-free survival

Di-ST - di-sialyl-T

DS - Dermatan sulfate

EBV - Epstein-Barr virus

ECM - Extracellular matrix

EGFR - Human epidermal growth factor receptor 1

EMT - Epithelial-to-mesenchymal transition

EXT1 - Exostosin 1

EXT2 – Exostosin 2

EXTL1 - Exostosin-Like 1

EXTL2 - Exostosin-Like 2

EXTL3 - Exostosin-Like 3

FGFR - Fibroblast growth receptor

FGF - Fibroblast growth factor

Fuc - Fucose

GAG - Glycosaminoglycan

Gal - Galactose

GalNAc - *N*-acetylgalactosamine

Glc - Glucose

GlcA – Glucuronic Acid

GLCE - D-Glucuronyl C5-Epimerase

GlcN - Glucosamine

GlcNAc - *N*-acetylglucosamine

GlcNAc6ST - N-acetylglucosaminyl-6-sulfotransferase

GlcNH3 - N-unsubstituted glucosamine

GPC – Glypican

GPI - Glycosyl-phosphatidyl-inositol

GS - Genomically Stable

HA - Hyaluronan

HER2 - Human epidermal growth factor receptor 2

HGF - Hepatocyte growth factors

HPSE 1 - Heparanase 2

HPSE 2 - Heparanase 2

HS - Heparan sulfate

HS2ST1 - 2-O-Sulfotransferase 1

HS3STs 1-6 - 3-*O*-Sulfotransferases 1-6

HS6ST1 1-3 - 6-*O*-Sulfotransferases 1-3

HSULF1 and 2 - HS 6-*O*-endosulfatase 1 and 2

IdoA - Iduronic acid

IGF-IR - Insulin-like growth factor 1 receptor

IgG - Immunoglobulin G

KO - Knock-out

KS - Keratan sulfate

KSGal6ST - KS galactosyl sulfotransferase

mAb - Monoclonal antibody

Man - Mannose

MAPK - Mitogen-activated protein kinase

MSI - Microsatellite instability

MW - Molecular weight

OE - Overexpression

OGA - *O*-GlcNAcase

OGT - *O*-GlcNAc transferase

OS - Overall survival

PAPS - 3'-phosphoadenosine-5'-phosphosulfate

PD-1 - Programmed cell death protein 1

PDGF - Platelet-derived growth factor

PGs - Proteoglycans

ppGalNAcTs - GalNAc-transferases

RON - Recepteur d'Origine Nantais

RTK - Receptor tyrosine kinase

SDC – Syndecan

Shh - Sonic Hedgehog

SLRPs Small leucine-rich proteoglycans

SPOCK - SPARC/Osteonectin CWCV and Kazal-like domain

ST - sialyl-T

ST6GalNAc1 - α2-6 sialyltransferase

TCGA-STAD - The Cancer Genome Atlas Stomach Adenocarcinoma

TGF-β - Transforming growth factor-β

TIDE - Tracking of Indels by Decomposition

TNM - Tumor-Node-Metastasis

UICC - Union for International Cancer Control

UST - Uronyl 2-O-sulfotransferase

VacA - Vacuolating cytotoxin A

VEGF - Vascular endothelial growth factor

VEGFR2 - Vascular endothelial growth factor receptor 2

Xyl – Xylose

XYLP - 2-Phosphoxylose phosphatase

XYLTs - O-Xylosyltransferases

β3GALT6 - β3-Galactosyltransferase 6

β3GAT3 - Glucuronyltransferase-I

β3GnTs - β-1,3-N-acetylglucosaminyl transferases

β4GalT - β-1,4-galactosyl transferase

Table of Contents

The glycan structures represented in this thesis follow the Symbol Nomenclature for Glycans (SNFG) format (1).

Chapter 1

General Introduction

General Aims and Objectives

This chapter includes text and figures adapted from the following literature review:

Marques C, Reis CA, Vivès RR, Magalhães A. Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression. *Front Oncol.* 2021 Nov 11;11:778752. doi: 10.3389/fonc.2021.778752.

1. General Introduction

1.1. Human Cancer

Cancer poses a great challenge in the 21st century, reflecting a profound impact in society, public health and economy at a global scale. It ranks amongst the top three leading causes of death for individuals aged 30-69 years in 177 of 183 countries, and in the year 2022 alone approximately 20 million new cases and 10 million cancer related deaths have been estimated (2). Cancer pathologies, also referred to as malignant tumors or neoplasia, are characterized by the abnormal and deregulated cell growth that often extends beyond the natural limits of the tissues from which these cells originate, ultimately invading adjacent tissues and organs. Cancer initiation and progression is driven by genetic alterations and a myriad of dynamic interactions between cancer cells, non-malignant cells, including immune, stromal and vascular cells, and matrix components, which might vary depending on the organs from which the tumor stems from (6). During this multistep process tumor cells must acquire specific oncogenic features to overcome many challenges and efficiently establish malignant tumors. These are termed cancer hallmark abilities, and include: sustaining proliferative signaling, evading growth suppressors and immune-derived destruction, avoiding cell death, enabling replicative immortality, inducing angiogenesis, invasion and metastasis, genome instability and mutation, deregulating cellular metabolism, acquisition of tumor inflammation (7). More recently, new emerging hallmarks are also being considered to be added to the list, including unlocking phenotypic plasticity, non-mutational epigenetic reprogramming, polymorphic microbiomes, and senescent cells (8).

1.2. Gastric Cancer

1.2.1. Epidemiology and Etiology

Gastric cancer currently ranks as the fifth most commonly diagnosed cancer and fifth leading cause of cancer related death worldwide (Fig. 1).

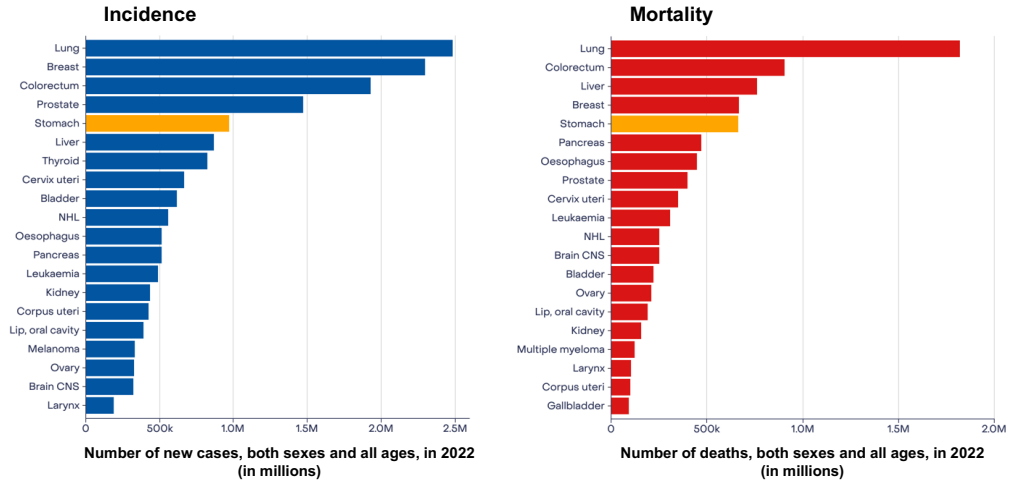


Figure 1. Number of new cases and deaths estimated for several types of cancer (world), for both sexes in all ages, in 2022. Graphic representation from GLOBOCAN 2022 (2).

In the year of 2022, approximately 1 million new cases were diagnosed and 660 000 people have died from this pathology (2). This classifies gastric cancer as an aggressive disease and the asymptomatic nature of the disease until later stages contributes to the high mortality to incidence ratio.

The geographic distribution of gastric cancer incidence is quite variable. Asian countries show higher incidence rates, followed by the Latin American and Caribean, and then European countries, whereas North American and African countries present the lowest incidence rates (Fig. 2) (2). The reason for these disparities is relatively unknown, although it has been associated with variations in the prevalence of certain risk factors (9). Incidence rates are also substantially higher in men, which are twice as likely to develop this disease as women (10). Gastric cancer is a multifactorial disease, influenced by both environmental and genetic factors. Besides gender, other risk factors are to be considered in the development of this malignancy, including age, ethnicity, smoking habits and dietary factors, such as high intake of nitrates, salt, smoked food and fat, low consumption of fresh fruits, vegetables and fiber (11). Host infection with pathogens like *Helicobacter pylori* (*H. pylori*) and the Epstein-Barr virus (EBV) has also been largely associated to gastric tumorigenesis. Particularly, infection with *H. pylori* is described as a main risk factor for gastric cancer (12).

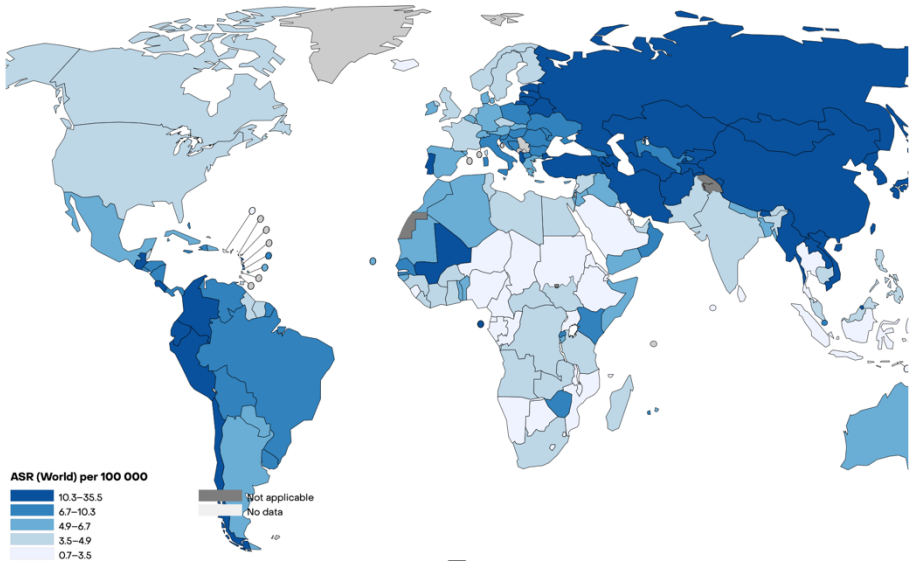


Figure 2. Age standardized incidence rates (world) of stomach cancer, for both sexes and all ages, in 2022. Heatmap from GLOBOCAN 2022 (2).

H. pylori, classified as a class I carcinogen, is a micro-aerophilic spiral-shaped gram-negative bacteria, capable of colonizing the host stomach epithelium from a very young age. It is resistant to the gastric highly acidic environment, and overtime it potentially induces chronic inflammation and gastritis. *H. pylori* infection classifies as one of the most common bacterial infections worldwide, affecting at least 50% of the world's population. Nonetheless most infected individuals remain asymptomatic, and only a smaller subset of

chronically infected individuals develop gastric cancer (13, 14). This outcome is potentially influenced by host genetic susceptibility, strain virulence and other environmental factors (15). The EBV is a ubiquitous infectious pathogen associated of 10% of all gastric cancer cases. It is still unclear the molecular mechanism through which EBV participates in cancer development (16).

Approximately 90% cases of gastric cancer are sporadic, while the remaining 10% are related to familial clustering, and subset of these (1-3%) are hereditary (17).

1.2.2. Classification and Pathogenesis

Over the years, various systems have been proposed for the efficient classification of gastric cancer based on histopathological features. The Lauren and the World Health Organization (WHO) classifications comprehend the two most commonly used (10).

The Lauren classification, established in 1965, divides gastric adenocarcinomas, which account for 90-95% of the gastric malignant cases, in two main histological subtypes: intestinal and diffuse (Fig.3) (18). A rarest and indeterminate subtype is also considered when the tumor tissue is undifferentiated to the point of lacking resemblance with either intestinal or diffuse subtypes, or mixed subtype if it presents histological features of both subtypes.

The intestinal subtype, which is the most frequent, is characterized by the presence of cohesive and moderately to well differentiated tumor cells and glandular structures. This subtype is intrinsically associated with environmental factors, and it comprehends a series of gradual events, usually over a long period of time, underlying intestinal subtype gastric malignant transformation. Each new event associates to a new histological profile (19). *H. pylori* infection plays a major role triggering this stepwise carcinogenesis pathway depending generally on the virulence of the infecting strain (higher invasion and colonizing abilities) but also on the host genetic susceptibility (inflammatory response). Persistent infection by highly virulent strains of *H. pylori,* positive for virulence factors cytotoxin-associated gene A (CagA) and the vacuolating cytotoxin A (VacA), can lead to aggravated chronic inflammation and gastritis, which are accompanied by progressive damage and structural changes in the gastric epithelium. Over a long period of time these lesions can progress to multifocal atrophic gastritis, associated to the loss of gastric glands and the appearance of fibrotic tissue, followed by intestinal metaplasia, where the gastric epithelium acquires an intestinal phenotype. It later can lead to dysplasia, a malignant pre-neoplastic lesion characterized by abnormal tissue architecture, and finally, to carcinoma (Fig. 3) (20).

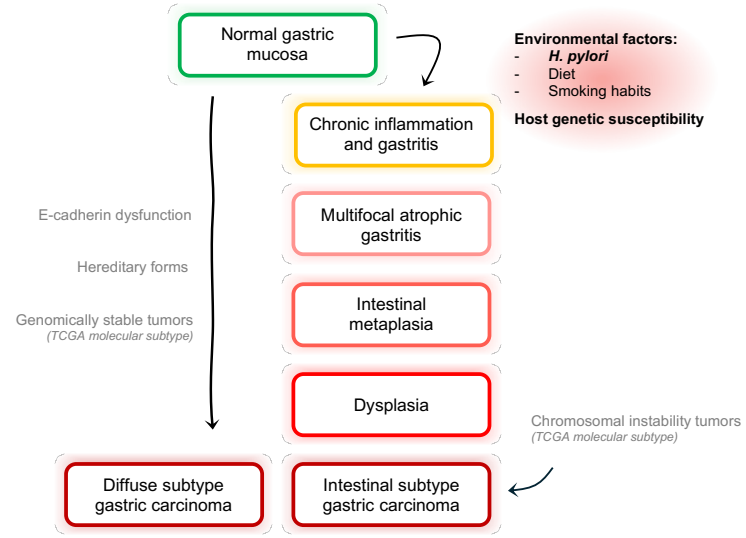


Figure 3. Gastric carcinogenesis main pathways associating with diffuse and intestinal subtypes. The intestinal and diffuse gastric carcinoma subtypes arise from different factors and induce distinct histological and molecular features. High exposure to specific environmental factors can induce stomach chronic inflammation and trigger the multi-step carcinogenesis cascade that often leads to intestinal gastric cancer, depending on the host genetic susceptibility. The diffuse subtype is mostly associated with spontaneous mutations, and often displays abnormal expression of E-cadherin. E-cadherin mutation and dysfunction is the main cause of the subtype hereditary forms. In regards to the TCGA molecular classification, most genomically stable tumors classify as diffuse, while chromosomal instable tumors are enriched in the intestinal histology.

The diffuse subtype accounts for a smaller but more aggressive percentage of cases. Diffuse carcinomas exhibit single, poorly cohesive/adhesive cells, with tendency to invade the mucosa and stomach wall, and lack well-structured glandular differentiation or gland formation. This specific subtype frequently includes the signet ring cell type and is mostly associated to familial clustering and hereditary forms. Moreover, it is greatly characterized by E-cadherin aberrant expression and/or deregulation, causing defective cell-cell adhesion and inducing a more invasive cell phenotype (Fig. 3) (21).

In 2014, The Cancer Genome Atlas (TCGA) Research Network proposed a new classification considering gastric cancer molecular features, and aiming to identify disrupted pathways and the potential promoters of the different classes of gastric adenocarcinomas (22). Four genomic subtypes have been highlighted within this new classification system: EBV-positive tumors, microsatellite instability (MSI) tumors, genomically stable (GS) tumors, and chromosomal instability (CIN) tumors. Within these different classes, the GS tumors, harboring mutations in E-cadherin (*CDH1*) and alterations in Rho-GTPases, are enriched in the histological diffuse subtype, while the CIN tumors, which present increased amplification of receptors tyrosine kinase (RTKs), have mostly an intestinal histology.

1.2.3. Diagnosis and Treatment

The asymptomatic nature of the disease up until later stages is a great medical concern and explains why it is still largely underdiagnosed. Indigestion, loss of appetite, weight loss

and abdominal pain are amongst the most common symptoms usually reported at later stages (23).

Endoscopic examination is the main diagnosis tool used in the clinical practice to determine the localization and macroscopic type of the tumor. Histological confirmation is then carried out through biopsy. Tumor staging assessment is fundamental to select the correct course of action for each patient, which could either be curative or palliative. This staging employs a Tumor-Node-Metastasis (TNM) system, which addresses tumor size (T stage), regional lymph node spread (N stage) and distant metastatic spread (M stage), and it is currently determined according to the American Joint Committee on Cancer (AJCC)/Union for International Cancer Control (UICC) TNM 8th edition staging manual (10, 23).

Treatment wise, endoscopic resection, for early gastric cancer, and surgical gastrectomy, to lymph node positive tumors from T1 to T4 stages, are the two main procedures in curative therapy. Preoperative, perioperative and postoperative chemotherapy and radiotherapy are additionally applied to improve the chances of curative resection and patients' overall survival, depending on the staging of the tumor and patients' health condition (10, 23).

The use of humanized monoclonal antibodies (mAb) for the targeted therapy of patients exhibiting genetic amplification or overexpression of specific molecular targets, has also significantly improved the clinical management of gastric cancer. Currently, there are three clinically approved targeted therapies, which include the use of Trastuzumab, Ramucirumab and Pembrolizumab antibodies to target the aberrant expression of human epidermal growth factor receptor 2 (HER2), vascular endothelial growth factor receptor 2 (VEGFR2) and the programmed cell death protein 1 (PD-1), respectively (24). There are several other targeted therapies that are currently under study and in clinical trials, aiming to target and inhibit the intercellular junction component Claudin-18.2, fibroblast growth factor receptor (FGFR) and human epidermal growth factor receptor 1 (EGFR) mediated pathways (25-28). Unfortunately, the emerging problematic of natural or acquired resistance of gastric cancer patients to the already implemented therapies in the clinical setting slightly dampened the therapeutic efficiency of these types of agents.

On top of the previously mentioned strategies of diagnosis and treatment, the use of serological markers has great potential as a predictive tool to help mitigate gastric cancer burden. There are currently three clinically approved serological assays used for detection of specific "glycobiomarkers" in gastric cancer patients, namely CA 19.9, CA 72.4, and CEA (29). The use of these markers is based on the scientific knowledge gathered over many years highlighting the role of glycosylation, and the appearance of aberrant glycan

structures, both in oncogenesis and as a driving factor in tumor pathophysiology (discussed in the following sections).

The glycoprotein antigen CA 19.9-based assay detects the glycan structure sialyl Lewis a, commonly present at high levels in gastric cancer. Due to the general ubiquity of this glycan, it is mostly used as a predictive factor for patient recurrence and response to therapy (30). The CA 72.4-based assay detects increased expression of the sialylated glycan sialyl-Tn antigen, whose serologic levels associate with tumor aggressiveness and patients' worse prognosis (31). Similarly to CA 19.9, high levels of this glycan antigen in premalignant lesions, restricts the use of this marker in diagnosis (32, 33). Lastly, the carcinoembryonic antigen (CEA)-based assay detects highly glycosylated CEA glycoproteins present at high levels in advanced gastric patients. Nonetheless, it holds sensitivity limitations, hampering early cancer detection (34, 35).

1.3. Protein Glycosylation

Glycosylation classifies as a highly abundant post-translational modification that takes places virtually in all mammalian cells. Through this modification, carbohydrate chains, commonly referred to as glycans, are added to different macromolecules, like lipids and proteins, leading to the formation of different classes of glycoconjugates that, together with other free polysaccharide chains, comprise the glycome (36). Cellular glycosylation occurs in the endoplasmic reticulum (ER) and Golgi apparatus, and involves the participation of a wide variety of glycosyltransferases, glycosidases and different glycan modifying enzymes, chaperones and transporters (36). This process is finely tuned at multiple levels and depends on numerous factors, including substrate availability, enzyme expression levels and activity, and enzyme localization within the secretory pathway (37, 38). Glycosylation is therefore frequently described as a "non-template" driven process, as it does not directly reflect gene expression.

The combination of different monosaccharides and maturation of the resulting polysaccharide chains lead to the formation of more complex glycan structures that, when covalently attached to either proteins or lipids, provide an additional layer of structural and functional heterogeneity. The different classes of glycoconjugates are divided according to the nature of the non-glycan portion of these macromolecules and the linkage of the sugar chain, and can be classified as glycosphingolipids, glycoproteins and proteoglycans (PGs). Recently a new class of glycoconjugates encompassing membrane bound glycoRNA has also been proposed (39). Focusing on protein glycosylation, this process entails the addition of at least one glycan chain to a polypeptide backbone, and it is defined by the variable structure and linkage of the added glycan chain(s). The different classes of protein-bound

glycans include *N*-glycans and *O*-glycans, as important components of glycoproteins, and glycosaminoglycans (GAGs), as main constituents of PGs (Fig. 4).

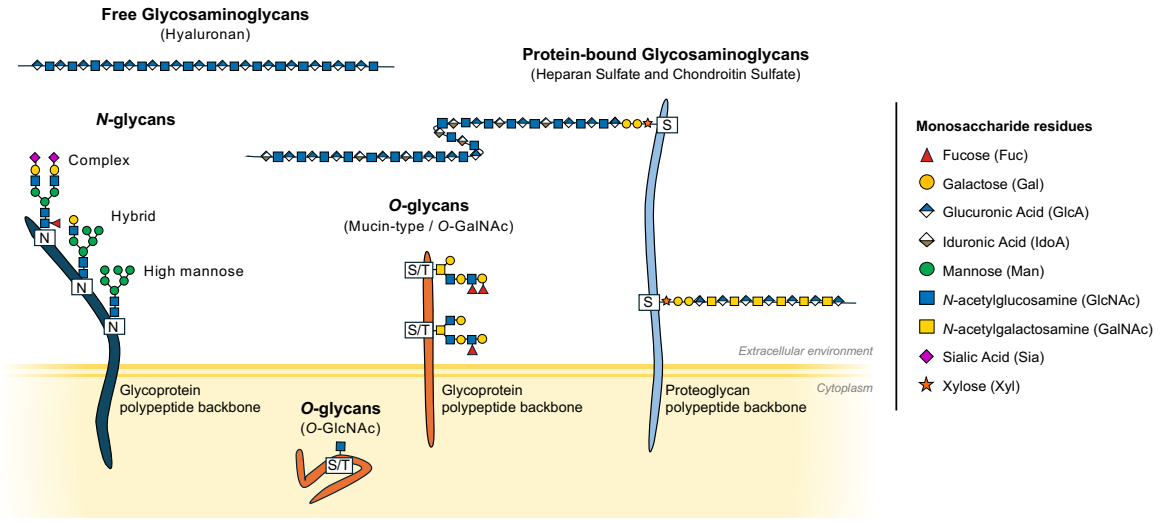


Figure 4. Main classes of free and protein-bound glycans, including *N*-glycans, *O*-glycans and glycosaminoglycans (GAGs), as important constituents of different glycoproteins and proteoglycans (PGs). Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

1.3.1. *N*-glycans

N-glycans pertain to branched oligosaccharide chains covalently linked to a nitrogen atom of asparagine (Asn) residues within a specific consensus sequence (Asn-X-Ser/Thr; where X can be any amino acid except proline (Pro)). All different *N*-linked glycans exhibit one similar feature: a defined pentasaccharide core region composed by two *N*-acetylglucosamine (GlcNAc) residues and three Mannose (Man) residues (Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn) (37). This core glycan results from extensive processing in the ER and Golgi, as a larger precursor glycan structure is initially added to the core protein (Fig. 5).

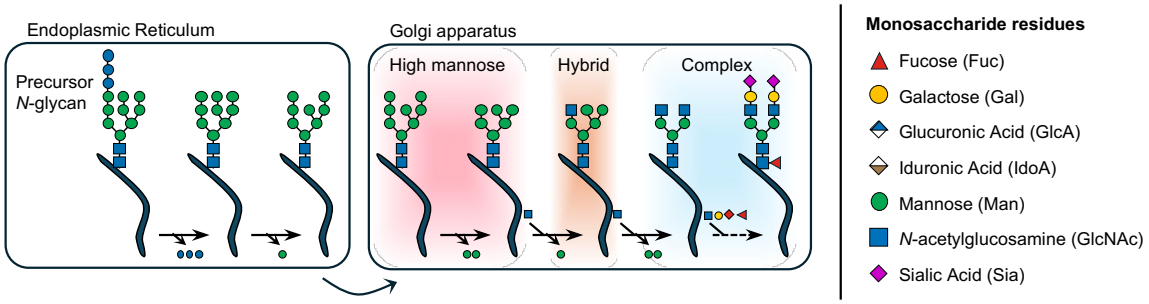


Figure 5. *N*-glycosylation pathway. Schematic representation depicting an abbreviated version of the main biosynthetic steps taking place in the endoplasmic reticulum (ER) and Golgi, leading to the production of high mannose, hybrid and complex *N*-glycans. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

The degree of processing of *N*-glycans precursor molecule gives rise to various structures i) high-mannose/oligomannose *N*-glycans, whose core is extended by Man residues (results from fewer processing reactions), ii) complex *N*-glycans, whose precursor

structure is a biantennary glycan (two branches) in which each Man is further elongated by the addition of at least one GlcNAc residue or iii) hybrid *N*-glycans, whose biantennary core is extended differently between the two branches - one of the branches is extended by Man residues (similar to high mannose *N*-glycans), while the other branch is extended by the attachment of an GlcNAc, that can be further elongated (similar to complex *N*-glycans). Hybrid and complex *N*-glycans can be further modified through galactosylation, GlcNAcylation, sialylation, fucosylation and sulfation reactions, contributing to the polymerization and modification of *N*-glycan branches and producing multiple mature and complex sugar structures (40).

1.3.2. *O*-glycans

O-glycans are covalently linked to serine (Ser) or threonine (Thr) hydroxyl groups within variable consensus domains, depending on the type of glycan being synthesized. There are multiple protein *O*-glycosylation pathways that lead to the formation of different classes of glycoconjugates. This process is initiated by the modification of the protein backbone with a specific monosaccharide residue, that can either be *N*-acetylgalactosamine (GalNAc), GlcNAc, Fucose (Fuc), Glucose (Glc), Galactose (Gal), Xylose (Xyl) or Man, followed by elongation into different structures (36). GalNAc-linked *O*-glycans are one of the most abundant classes of polysaccharides present in a myriad of membrane bound and secreted glycoproteins (in fact, virtually all proteins are susceptible to this modification). These glycans are also referred to as mucin-type *O*-glycans, due to their presence as dense saccharide aggregates in mucins. Mucins are large glycoproteins that feature repeating domains rich in Ser and Thr residues, and therefore are more prone to undergo extensive modification with several GalNAc-linked *O*-glycans (37, 41).

GalNAc-linked *O*-glycans biosynthesis is initiated in the Golgi by polypeptide GalNAc-transferases (ppGalNAcTs), that catalyze the addition of one GalNAc residue to a Ser/Thr residue, forming the Tn antigen (GalNAcα1–*O*-Ser/Thr) (42). This structure will be further targeted by other Golgi transmembrane glycosyltransferases as it moves from *cis*-Golgi to *trans*-Golgi compartments. During this pathway, the Tn antigen can be sialylated by the α2-6 sialyltransferase (ST6GalNAc1), forming the short *O*-glycan sialyl-Tn antigen. Otherwise, it can be extended and converted into 4 different types of core structures, that can be differently elongated at later stages (Fig. 6). The core 1 structure (Galβ1-3GalNAcα1-*O*-Ser/Thr), also known as T antigen, is the most abundant core structure, and it is generated by the addition of a Gal residue to the Tn antigen by a Gal-transferase (C1GalT1). Further modification of this core 1 glycan by core 2 β1-6 *N*-acetylglucosaminyltransferases (C2GnTs), which catalyze the addition a GlcNAc branch, results in the formation of the core 2 structure (GlcNAcβ1-6(Galβ1-3)GalNAcα1–*O*-Ser/Thr). In addition, core 1 structure,

similar to the Tn antigen, might be capped with a sialic acid to form the sialyl-T (ST) or di-sialyl-T (Di-ST) antigens. This specific reaction is catalyzed by the α2-3 sialyltransferase (ST3Gal1). Alternatively, the Tn antigen can be targeted by the core 3 β1-3 N-acetylglucosaminyltransferase 6 (B3GNT6) and undergo extension with a GlcNAc residue to form the core 3 glycan (GlcNAcβ1-3GalNAcα1–*O*-Ser/Thr). Following, core 3 branching with a GlcNAc residue, by the C2GnNT3, prompts the formation of a core 4 structure (GlcNAcβ1-6(GlcNAcβ1-3)GalNAcα1–*O*-Ser/Thr) (40). These core structures are frequently elongated by distinct glycosyltransferases, sulfated and capped with sialic acid and fucose moieties at later stages, dictating *O*-glycans synthesis termination (Fig. 6).

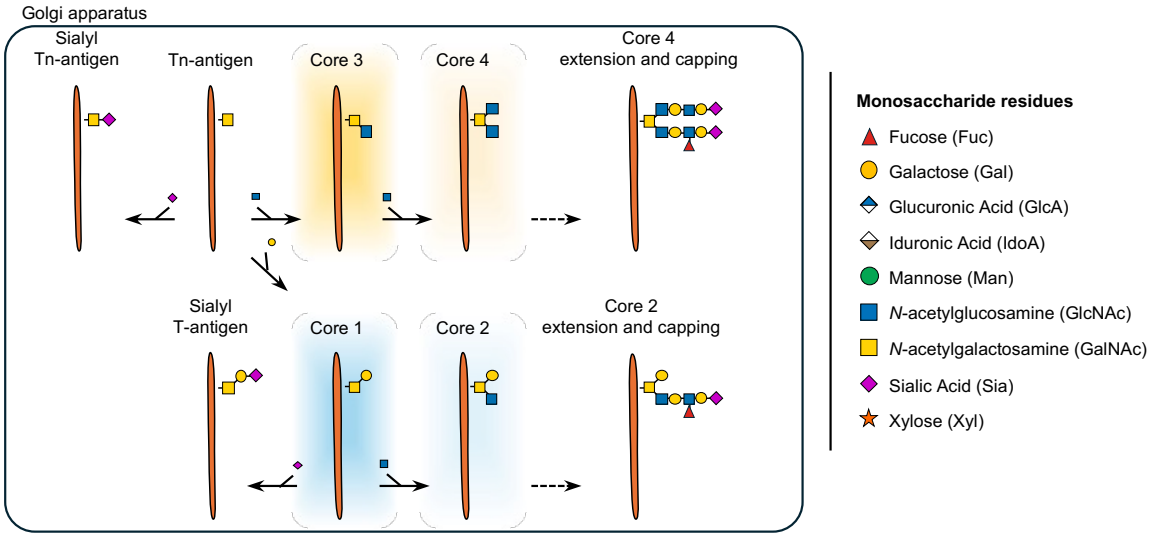


Figure 6. *O*-glycosylation pathway. Schematic representation depicting an abbreviated version of the main biosynthetic steps taking place in the Golgi, that lead to the production of the distinct *O*-glycan structures. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

Contrary to what is described for GalNAc-linked *O*-glycans, the biosynthesis of other types of *O*-glycans before mentioned, including *O*-Fuc, *O*-Glc, *O*-Gal, *O*-GlcNAc, *O*-Man and *O*-Xyl, is initiated in the ER (36). Additionally, *O*-GlcNAcylation targets specifically cytosolic and nuclear proteins, in a highly dynamic process catalyzed by *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA), that promptly add or remove the GlcNAc residues from the targeted proteins. Over the years, many studies have highlighted a strong and complex interplay between this process and proteins phosphorylation (43).

1.3.3. Glycosaminoglycans

GAGs are long and linear polysaccharides with a defined repeating disaccharide sequence. These glycan chains exhibit remarkable structural features that set them apart from the previously mentioned polysaccharides, including their extraordinary length that can range up to hundreds of disaccharide units. Due to the elongated structure of these glycans, proteins GAGosylation (process through which proteins are modified by a GAG chain)

account for a significant portion of the resulting glycoconjugates, known as PGs, influencing both structural integrity and cellular function. Additionally, GAGs can be classified into five different classes, according to their repeating disaccharide units: hyaluronan (HA), heparin/heparan sulfate (HS), chondroitin sulfate (CS), dermatan sulfate (DS) and keratan sulfate (KS) (Fig. 7) (4).

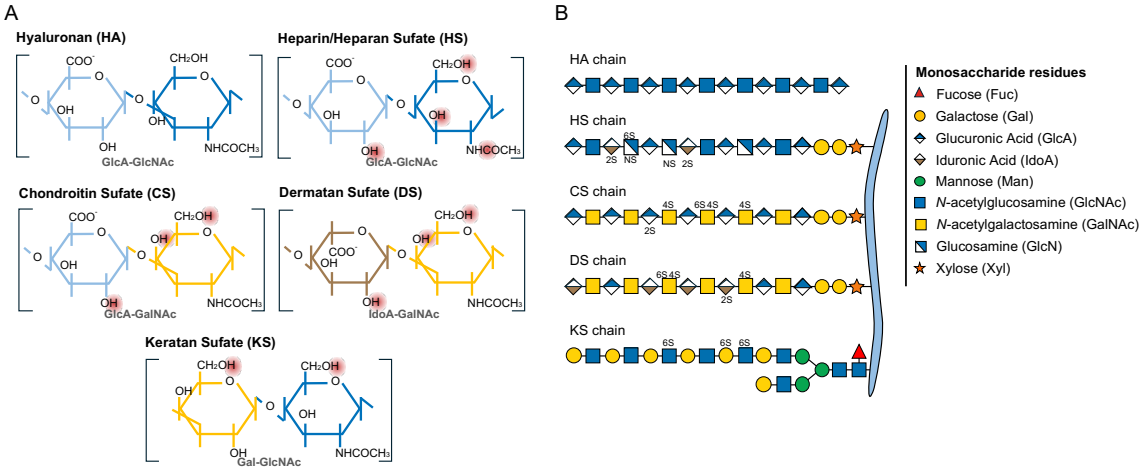


Figure 7. Glycosaminoglycans (GAGs). (A) GAGs are divided in 5 major classes according to their composing disaccharide unit repeats: hyaluronan (HA), heparin/heparan sulfate (HS), chondroitin sulfate (CS), dermatan sulfate (DS) and keratan sulfate (KS). **(A** and **B)** Different GAG disaccharides can be sulfated at multiple positions, except for the non-sulfated and protein-free HA. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

HA stands out as a unique GAG, particularly known for its elongated, and therefore high-molecular weight, and non-sulfated or modified structure. Composition wise, it is formed by repeating disaccharide units of GlcNAc and glucuronic acid (GlcA) (GlcNAcβ1-4GlcAβ1-3), and unlike the other GAGs, it is not covalently attached to a protein moiety. In eukaryotic cells, HA is polymerized as a free glycan at the inner surface of cells plasma membrane by the transmembrane enzymes hyaluronan synthases (HAS 1-3) and immediately pushed out to the cell surface or extracellular matrix (ECM) (44).

The remaining classes of GAGs are sulfated and bound to proteins, existing as important components of multiple PGs, and their biosynthesis occurs in the ER and Golgi.

HS and CS chains, although with different disaccharide composition, are both covalently attached to Ser residues of the protein moieties within specific Ser-Gly sequons, which are known to be in close proximity of two acidic amino acid residues. The flanking regions of these GAG consensus domains might vary and, in turn, influence the type of GAG attached to the protein backbone (45). Both classes of GAGs consist of repeating disaccharide units formed by GlcA residues that alternate with either GlcNAc to form HS (GlcNAcα1-4GlcAβ1-4), or GalNAc to form CS (GalNAcβ1-4GlcAβ1-3).

HS and CS share the initial step of their biosynthetic pathways: the formation of a tetrasaccharide linker, assembled by one Xyl, two Gal and one GlcA residue, through which

both types of chains are covalently attached to their respective core proteins. This is followed by chain elongation promoted by either HS- or CS-specific glycosyltransferases that will add the alternate sugar residues that form each GAG polysaccharide sequence. Finally, still during polymerization, GAG chains will be subjected to several modification reactions, by different HS or CS modifying enzymes, including epimerization of GlcA into iduronic acid (IdoA), and several *N*- and *O*-sulfation reactions that confer high structural variability, and consequently function diversity, to these GAGs. Nomenclature wise, if any GlcA residue that constitutes CS chains is epimerized to IdoA, this GAG is then named DS (also known as CS B) (46, 47). The work outlined and discussed in this thesis was mainly centered on GAGs, specifically focusing on HS and, to a lesser extent, CS. Additional detailed information on these subjects will be elaborated in greater detail in the following sub-section of this chapter.

Lastly, KS chains biosynthesis varies depending on the glycosidic bond it shares with the respective PG, which also influences the structure of the resulting GAG chain. KS can either be *N*-linked, through the covalent attachment of a GlcNAc to an Asn, to form KS type I, or it can be *O*-linked to a Ser/Thr residue, *via* a GalNAc, to form the KS type II. Additionally, it has also been described a KS type III chain, whose biosynthesis is initiated with the addition of a *O*-linked Man residue to a Ser amino acid (48). A distinctive feature from these carbohydrates is the presence of a sulfated poly-*N*-acetyllactosamine chain, composed by repeating disaccharides containing GlcNAc and Gal residues (Galβ1-4GlcNAcβ1-3), which are also found in glycoproteins. The elongation of KS chains into this structure is catalyzed by the β-1,4-galactosyl transferase 1 (β4GalT-1) and β4GalT-4, and the β-1,3-*N*-acetylglucosaminyl transferases 1-7 (β3GnT 1-7). The polymerized KS chains are also subjected to several modification reactions, known to influence KS biological roles. KS sulfation is described as the main modification, and it is mainly conducted by *N*-acetylglucosaminyl-6-sulfotransferase (GlcNAc6ST) and KS galactosyl sulfotransferase (KSGal6ST), which sulfate GlcNAc and Gal residues, respectively. Additional sialylation and fucosylation reactions can also take place (49).

1.3.4. Gastric cancer-associated glycan expression

Protein glycosylation plays a crucial role in numerous physiological events, endowing proteins with additional biological functions. These functions include structural modulation, self-recognition and communication, and management of extrinsic interactions (50). At the molecular level, glycosylation is critical in controlling proteins folding and stability, and conferring protection against proteolysis. A myriad of cell surface glycans, also known as the glycocalyx, play significant roles in cell motility, cell-cell and cell-ECM adhesion, signaling and immune recognition. Within the ECM, varying compositions of different

carbohydrates influence structural organization, while also being capable of binding and determining the gradient of many soluble molecules. At the tissue level, glycans serve as a physical protective barrier against microorganisms' invasion. On the opposite hand, specific glycans might be recognized and targeted by pathogens, mediating invasion and colonization (e.g. *H. pylori* recognizes specific glycan structures in the human gastrointestinal tract mucosa, which influence pathogenesis) (50). Overall, these and many other specific roles taken by glycans, implicate their participation in many biological events that are inherent to cell and tissue growth and development and in organisms' homeostasis. However, deregulation of glycans biosynthetic pathways is frequently described as a driving feature in several different disorders, and it often contributes to biochemical and biophysical changes, influencing disease onset, progression and ultimately poor outcomes.

In the context of cancer, altered changes in glycosylation are associated to malignant transformation and disease progression. Over the years, a lot of research has been dedicated to study cancer-associated glycans, more specifically the aberrant glycosylation patterns associated with different types of cancer and, within the same pathology, to different stages of the disease (from pre-malignant lesions to advanced stages of carcinogenesis). Additionally, there has been a growing interest in learning the events that originate these changes, as well as understanding the advantages that these altered glycans provide to cell malignant transformation and to tumor cell signaling and behavior (51, 52). Aberrant glycosylation in cancer can arise from altered expression or localization of specific glycosyltransferases, glycosidases or chaperones and changes in the availability of either the protein or sugar substrates, amongst other (53). Such changes can drive tumor progression and dissemination by influencing cancer cell signaling cascades, cell proliferation and death evasion, invasion and metastasis, angiogenesis and immune escape (52, 54-56). The disclosure of new glycan epitopes and their mechanisms of action in cancer-related events holds significant value in the clinical setting, as these molecules can potentially function either as important biomarkers for early disease detection (57), or as targets in tumor cells for selective and efficient drug delivery in cancer therapy (38, 58).

In the context of gastric cancer, there are major aberrant glycan features that have been detected in pre-malignant lesions, and throughout the different stages of the disease, accompanying cell malignant transformation and main oncogenic events (59). These include the expression of truncated *O*-glycans, increased *N*-glycan branching and increased terminal sialylation and fucosylation.

O-glycan truncation, resulting from incomplete elongation of *O*-glycans, leads to the accumulation of T (core 1) and Tn antigens, and their sialylated counterparts ST and STn antigens. It can be driven by different mechanisms including altered expression of glycosyltransferases, like ST6GalNAc1 (60, 61), glycosyltransferases' mislocalization (62)

and downregulation of the COSMC molecular chaperone, an important player in the synthesis of core 1 structures (63). These glycan structures are rarely detected in normal tissues, but highly expressed in tumors, including gastric cancer (60, 64). Studies showed that in this latter, increased levels of shorter *O*-glycans change the activation state of important RTKs, including EGFR and HER2, and promote cell migration and invasion (65, 66).

Increased *N*-glycan branching identified in gastric cancer was associated to upregulation of the branching glycosyltransferase *N*-acetylglucosaminyltransferase-V (GnT-V or MGAT5) and downregulation of the bisecting glycosyltransferase GnT-III (as this enzyme synthesizes a glycan substrate that cannot be further branched by GnT-V). These highly branched structures are known to impact integrin and E-cadherin functions, and consequently cell-cell and cell-ECM interactions, and have been implicated in cancer patients' poor prognosis (56, 67-69).

Abnormal overexpression of the α2-6 sialyltransferase (ST6Gal1), an enzyme that regulates the terminal addition of sialic acids to branched *N*-glycans, was also found to have a negative impact in multiple malignancies, including gastric cancer (70). In addition, α2-6 sialylation of *N*-glycans that are commonly linked to the members of the EGFR family, EGFR (71) and HER2 (72), was also revealed to modulate receptor activation and to contribute to the resistance of gastric cancer cells to anti-HER2 trastuzumab-based therapy (73, 74).

Another major distinctive glycosylation feature found in gastric cancer relates to the presence of high levels of terminal sialylated glycans sialyl Lewis x and sialyl Lewis a. Altered expression of glycosyltransferases underlies the formation of these aberrant structures in many different types of malignancies, including gastric cancer, and it correlates with patients' poor survival (33, 75, 76). Furthermore, RTKs modification with these specific sialylated glycans, such as r*ecepteur d'origine nantais* (RON) and hepatocyte growth factor receptor (c-Met), often leads to increased activation and has been highlighted as a possible mechanism underlying gastric cancer cell aggressive phenotype (77, 78).

Regarding GAGs, these represent a sub-class of molecules whose impact in the oncobiology field appeared to progress at a slightly slower pace. GAGs exhibit an incredible complex structure, from the outstanding length of most types of chains, to the highly variable sulfation and epimerization profiles, not only within the same GAG chain, but also over different tissues. These features required the development of new sophisticated technology over the years. Even nowadays, multiple analytical tools might have to be employed simultaneously to reveal new information on structural disaccharide features and sequence, GAG-protein interactions and overall GAG functional roles in maintaining homeostasis and in disease (79). Nevertheless, significant progress has been made, and although a clear general cancer related GAG aberrant feature is still yet to be identified, numerous studies

have shown that GAGs and PGs are key players in a myriad of different oncogenic events, further highlighting the clinical potential of such molecules in cancer therapy. On top of that, in recent studies specific GAG profiles have been identified as possible new predictive tools (80-82). Specifically in the context of gastric cancer, it has been hard to describe a pattern in GAGs and PGs expression and distribution, but recent scientific evidence shows that HS-related genes and HSPGs might play a key role in tumor cell biology (to be further discussed in sub-section 1.7.3).

1.4. Heparan Sulfate Proteoglycans

HS chains are long, linear and highly anionic polysaccharides covalently attached to specific protein moieties, known as HSPGs. These glycoconjugates are ubiquitously expressed on cell glycocalyx and ECM, as well as in secreted extracellular vesicles (EVs) (83-85), and control several regulatory mechanisms including the formation of extracellular gradients, cellular growth and proliferation, cell adhesion, migration and invasion, membrane trafficking and angiogenesis (86, 87). By interfering with these cellular events, HSPGs display important functions both in physiology and pathology, controlling embryonic development, ECM assembly and maintenance, tissue remodeling, metabolism homeostasis, pathogen invasion and inflammation (5, 88-93). Particularly in cancer, HSPGs and HS chains hold very relevant roles in the development and progression of the disease. Interaction with different ligands and structural proteins, and modulation of several signaling networks, prompts HSPGs active participation in cell transformation and proliferation, tumor growth and metastasis, amongst other cell oncogenic events (5, 84, 94), highlighting the fundamental need of further studying HS and HSPGs biosynthesis in the context of cancer.

Close to 80 different PG core proteins have been reported up until now, some of which are described as “part-time” PGs, with variable degrees of occupancy of their composing GAG sites (95). Structure wise, PGs are extremely heterogeneous, exhibiting significant differences in terms of protein sequence and structure, and in the number and nature of their associated GAG chains. As described by Iozzo and Schaefer, these glycoconjugates can be classified into four groups according to their cellular and subcellular localization, gene/protein homology and the shared presence of “unique protein modules” (which refers to specific function units within protein moieties with biological impact): (i) intracellular, (ii) cell surface, (iii) pericellular, and (iv) extracellular PGs (Fig. 8A) (3).

A

LOCATION / CLASSIFICATION			PROTEOGLYCANS / GENE	GAG TYPE AND NUMBER	CORE PROTEIN MW (kDa)
Intracellular	Secretory granules		Serglycin / *SRGN*	10-15 Hep/CS	10-19
Cell surface	Transmembrane		Syndecan 1-4 / *SDC 1-4*	1-3 HS ; 0-2 CS	31-45
			NG2 / *CSPG4*	2-3 CS/DS	251
			Betaglycan / *TGFRB3*	0-1 HS; 0-1 CS	110
			Phosphacan / *PTPRZ1*	2-5 CS/DS	175
	GPI-anchored		Glypican 1-6 / *GPC 1-6*	1-3 HS	60
Pericellular	Basement membrane		Perlecan / *HSPG2*	1-3 HS; 0-2 CS	400
			Agrin / *AGRN*	1-3 HS	200
			Collagen type XVIII / *COL18A1*	2-3 HS	147
			Collagen type XV / *COL15A1*	CS/HS	142
Extracellular	Hyalectan Lectican		Aggrecan / *ACAN*	~100 CS/ ~20 KS II	208-220
			Versican / *VCAN*	0-15 CS/DS	265
			Neurocan / *NCAN*	1-2 CS/DS	145
			Brevican / *BCAN*	0-4 CS/DS	96
	SRLPs	Class I	Biglycan / *BGN*	0-2 CS/DS	38
			Decorin / *DCN*	1 CS/DS	36
			Asporin / *ASPN*		43
			ECM2 / *ECM2*		77
			ECMX / *ECMX*		
		Class II	Fibromodulin / *FMOD*	2-4 KS I	59
			Lumican / *LUM*	3-4 KS I	37
			PRELP / *PRELP*	KS	44
			Keratocan / *KERA*	3-4 KS I	37
			Osteoadherin / *OMD*	KS	49
		Class III	Epiphycan / *EPYC*	DS/CS	37
			Opticin / *OPTC*		37
			Osteoglycin / *OGN*	2-3 KS I	25
		Class IV	Chondroadherin / *CHAD*		41
			Nyctalopin / *NYX*		52
			Tsukushi / *TSKU*		38
		Class V	Podocan / *PODN*		67
			Podocan-Like 1 / *PODNL1*		57
		SPOCK	Testican, 1-3 / *SPOCK 1-3*	HS	47-49

B

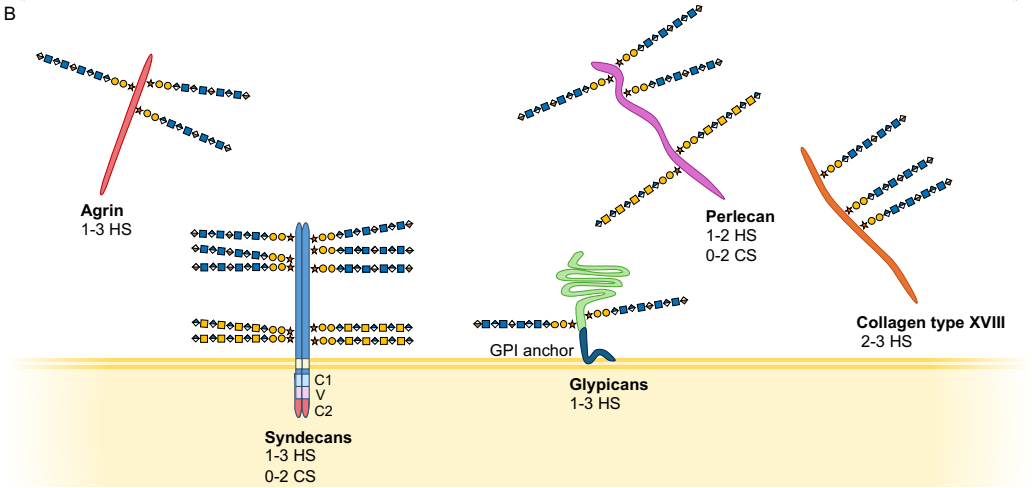


Figure 8. Proteoglycans (PGs) classification. (A) Table including the distinct groups of PGs classified according to their cellular and subcellular localization, gene/protein homology and function. Adapted from (3). Data on the number of glycosaminoglycans (GAGs) and core protein MW was also retrieved from (4) and UNIPROT. **(B)** Illustrative representation depicting the main families of cell surface and extracellular matrix (ECM) Heparan Sulfate PGs (HSPGs). The number and type of GAG chains associated with each PG type was also included in the figure. Adapted from (5).

1.4.1. Intracellular PGs

Serglycin stands as the only intracellular PG and the only one that is modified with heparin chains, although it can also be modified with HS and CS. Serglycin core protein is enriched in Ser and Gly residues, which comprise its many clustered GAG attachment sites, and the GAGosylation of this molecule depends on the cell in which it is expressed and it influences its functional roles (96). Serglycin is expressed in hematopoietic cells, including connective tissue mast cells, where it is preferentially modified with heparin, and other

inflammatory cells like lymphocytes, neutrophils, eosinophils and platelets, where it carries mostly CS chains. HS modified Serglycin has also been reported in mouse macrophages (96). This PG is able to bind to and regulate the storage and release of immunoactive components, including chemokines, cytokines, growth factors and proteases, thus participating in the inflammatory process (97). Although this cannot be exclusively attributed to the presence of the GAG chains, it has been shown that highly anionic and sulfated heparin chains are crucial to Serglycin's storage function in secretory granules within connective tissue mast cells, and therefore hypothesized that it relies on GAG-protein electrostatic interactions (98). Serglycin can also be expressed in endothelial cells and chondrocytes where it has been implicated in chemokine and protease secretion (99, 100)

1.4.2. Cell surface PGs

Cell surface PGs can be further divided in two sub-classes depending on how they are bound to the cell membrane: the transmembrane PGs, which are directly bound through their protein core, and the glycosyl-phosphatidyl-inositol (GPI)-anchored PGs. The sub-class of transmembrane PGs includes the family of Syndecans 1-4 (SDCs 1-4), NG2, Betaglycan, Phosphacan, CD44, CD74 and thrombomodulin, while the GPI-anchored PGs are constituted by the Glypicans 1-6 (GPCs 1-6). A great variety of PGs within this class are HS-carriers, and are generally known to participate in multiple cell signaling cascades as important co-receptors, interacting with different growth-factors and modulating the interactions of these soluble ligands with their targeted receptors. SDCs and GPCs are distinguished as the two major families of cell surface HSPGs (Fig. 8B).

The four members of the SDC family – SDCs 1-4 – are type I transmembrane proteins formed by an extracellular N-terminal domain, a transmembrane domain and a cytoplasmic C-terminal domain. This later domain has two conserved regions (C1 and C2) separated by a variable region (V) specific to each SDC, but conserved between species. The transmembrane domain includes specific motifs, including GXXXG and GXXXA, important for SDCs self-assembly and formation of stable homo- and hetero-dimers, an important structural feature implied in SDC functions (101, 102). Both the transmembrane and cytoplasmic domains are largely conserved across different SDCs, unlike the extracellular domain. The extracellular domain exhibits low sequence homology among all SDC members, except for the GAG attachment sites located within this region, typically at the opposite extremity of the cell surface (103, 104). SDCs core protein molecular weight (MW) ranges between approximately 30-45 kDa and their expression is cell and tissue dependent and can vary during development and tissue reorganization. Generally, SDC1 is expressed mostly in epithelial and plasma cells, SDC2 is expressed in mesenchymal cells and SDC3 is mainly present in neural tissues and developing musculoskeletal tissues. Unlike these

members, SDC4 expression is ubiquitous, although at lower levels (105). SDCs also differ from each other in terms of GAG content. SDC1 and SDC3 are described as hybrid molecules that are modified with both HS, localized towards the N-terminal domain, and CS chains, which are proximal to the plasma membrane. On the other hand, SDC2 and SDC4 are main HS carriers. Occasionally, certain SDCs, like SDC4, can be alternatively modified with CS chains (106, 107).

Mammalian SDCs display a wide range of biological functions, mostly through their ability to bind multiple biological active molecules. SDCs biological roles depend both on their cytoplasmic and extracellular domain. Many SDC-protein interactions established at the cell surface or ECM (upon protease cleavage) are mediated by the GAG chains of SDCs and a few rely on SDCs core protein. At SDC extracellular domain, GAG chains interact directly with multiple growth factors, including vascular endothelial growth factors (VEGF) and fibroblast growth factors (FGFs) to participate in vascular development and repair, cell differentiation and proliferation (108, 109), and ECM molecules, namely fibronectin, vitronectin and collagens, allowing them to mediate cell adhesion. Additionally, they can also bind several cytokines, chemokines and enzymes, overall modulating several functions related to cellular growth, differentiation and motility (105, 110). SDCs are also involved in the formation (111), and uptake of EVs acting as internalizing receptors expressed by the recipient cells (112). SDCs core protein can also bind directly to integrins and cell surface receptors, including EGFR and HER2, and mediate cell attachment and spreading, migration and survival (113). Additionally, SDCs extracellular protein moiety can be targeted by metalloproteinases and undergo a process called ectodomain shedding. This process produces soluble SDC fragments that are released into the extracellular environment, and ultimately regulates the levels of cell surface HS and the extracellular gradient of HS-binding ligands (114). On the other hand, SDCs cytoplasmic domains are known to establish interactions with intracellular kinases and actin, eliciting changes in cell cytoskeleton (115).

The mammalian GPC family is composed by six members – GPCs 1-6 – attached to the outer surface off cellular membranes through a C-terminal GPI anchor. The protein moiety of these molecules weights between 60-70 kDa. Although knowledge on GPCs protein structural features is still limited, it is known that they include a short hydrophobic C-terminal domain with a signal sequence for the attachment of the GPI anchor, followed by a larger tightly packed α-helice domain with a globular shape towards the extracellular side (N-terminal domain). Despite GPC members low homology, they all share the localization of fourteen cysteine residues at the globular shaped domain, through which seven disulfide bounds are formed (116, 117).

GPCs can carry up to four HS chains and, unlike SDCs, GPCs' HS attachment sites are localized closer to the C-terminal domain, immediately before the larger domain, in an

orientation that allows ligands and receptors proper interaction with both GAGs and protein core. This class of HSPGs is mostly involved in the modulation of morphogen signaling pathways driven by Wnt, the Bone Morphogenetic Proteins (BMP) and Sonic Hedgehog (Shh) (118, 119). Interestingly, different members of the GPC family have been shown to be able to both stimulate or inhibit these signaling pathways, *via* different mechanisms of action (116). GPCs can also be released to the ECM, both through proteolytic and phosphatidyl-inositol cleavage by proteases and lipases, respectively, participating in morphogens gradient formation (120).

1.4.3. Pericellular PGs

Within the class of Pericellular PGs are represented the GAG-modified glycoconjugates that are present in the basement membrane and in the pericellular space. These are high-molecular PGs, mostly modified by HS chains, that are anchored *via* integrins or other receptors (3). Perlecan, Agrin and Collagen XVIII are the main HS-modified constituting members of this class, playing distinct roles in homeostasis, including basement membrane maintenance, growth factor binding and modulation and angiogenesis (Fig. 8B) (121). Collagen XV, which is also a member of this class, is modified with CS.

Perlecan is a large HSPG broadly expressed in vascular and avascular tissues, whose protein core weights close to 400 kDa and encompasses 5 different domains (I-V). Each domain associates with different ligand partners and functions, eliciting the participation of this macromolecule in multiple physiological events (121, 122). The domain I (N-terminal) includes three HS attachment sites known to be important to FGF, VEGF, platelet-derived growth factor (PDGF), BMPs interactions. The domain II interacts with lipids, Wnt and Shh morphogens, participating in gradient formation, the domain III binds FGFs, and the domain IV, which is the largest, exhibits cell attachment and matrix stabilizing properties, by binding to structural matrix components like nidogen, fibronectin, and fibulin-2. The domain V (C-terminal; proteolytically cleaved endorepellin) shares interactions with multiple growth factors, integrin, and ECM molecules, and can evoke both pro- and antiangiogenic events (123, 124). Overall, Perlecan participates in cellular proliferation, differentiation and tissue development, cell adhesion, wound healing and inflammation, amongst many other cell events.

Agrin is another abundant and large HS carrier, with a size of 225 kDa, containing at least three HS attachment sites. The most relevant mammalian function attributed to this HSPG is related to the regulation and formation of neuromuscular junctions (125).

Collagen XVIII is localized in epithelial and vascular basement membranes and contains three HS attachment sites. Structure wise, it forms an homotrimer of three identical α1 chains, and its constituting N-terminal and C-terminal domains are spaced out by ten

interrupted collagenous domains. Collagen XVIII is mostly known for its role as a negative regulator of angiogenesis (3, 126).

1.4.4. Extracellular PGs

Finally, the extracellular PGs englobe the largest class of PGs, mostly modified by CS/DS and KS. These molecules can be divided in the subclass of Hyalectan letican, the Small leucine-rich proteoglycans (SLRPs) and the SPOCK (SPARC/Osteonectin CWCV and Kazal-like domain) family. The Hyalectan subclass is formed by four hyaluronan and lectin binding PGs: Agrecan, Versican, Neurocan and Brevican. These elements share important structural features, namely the presence of a hyaluronan-binding N-terminal domain, a central domain modified with GAGs, and a lectin binding C-terminal domain (127). By binding to HA, Hyalectan members display important structural roles in cartilage, blood vessels and central nervous system. The SLRPs subclass englobes at least 18 different members, formed by a relatively small core protein (~30-40kDa) consisting of leucine-rich repeats, that are ubiquitously expressed in most ECMs of all tissues. These PGs are known to bind multiple collagens, growth factors and cell surface receptors, and to participate in embryonic development and tissues homeostasis (128). SRLPs are grouped in five classes, the classes I, II, and III comprehend the canonical genes, and the classes IV and V englobe the non-canonical genes, amongst which are a few members that lack GAG chains but share large structural homology and functional similarities with the other members (3). Lastly, the three testican members (1-3) that form the SPOCK subclass are mostly described as calcium binding HSPGs that participate in neuronal regulation and are present in the testis and central nervous system (3).

In addition, there are a few extracellular PGs that can be in circulation instead of being localized in the ECM, such as endocan. Endocan is a small “part-time” PG modified with a relatively short DS chain (~32 disaccharides) secreted by endothelial cells. Its soluble form can circulate in the blood stream and exert regulatory functions at distant sites. Endocan DS chains can bind hepatocyte growth factor (HGF) and possibly modulate the activity of the Met receptor and downstream signaling cascades (129). Endocan can also bind to lymphocytes and monocytes and modulate its tissue adhesion (130), and stimulate endothelial cells VEGF-A/C mediated proliferation and migration (131).

1.5. Heparan Sulfate Biosynthesis

PG core proteins are of crucial importance to determine PGs localization, the associated GAGs and, in certain cases, to directly bind ligands and transduce signal (105, 110, 120). However, as seen by many of the beforementioned HSPGs, these molecules exert a vast majority of their known cellular functions through their HS chains, which.

depending on their structural features, can interact non-covalently with numerous proteins, commonly referred to as HS-binding proteins (132).

Structure wise, HS native chains are composed by repeating GlcNAcα1-4GlcAβ1-4 disaccharide units that, during and post biosynthesis, can suffer multiple modification reactions. This biosynthetic pathway occurs at the ER-Golgi apparatus interface and in the Golgi apparatus, and includes: i) assembly of a GAG-protein linker, which initiates the covalent binding of HS to PG core proteins; ii) the polymerization of the HS chain; and iii) the structural modification of the elongated chain (90). The first two stages of HS biosynthesis involve the stepwise sequential transfer of sugar residues to the growing chain and are catalyzed by different glycosyltransferases. The polymerizing chain then undergoes maturation by several HS modifying enzymes, both during and after biosynthesis.

The biosynthesis of GAGs, and HS in particular, is not a template-driven process, meaning that the extent of modification reactions that the glycan chains may undergo, and the consequent HS final structural motifs, are not directly encoded in the genome. The current general view is that HS biosynthesis is regulated by many factors, including the expression and availability of the enzymes involved in this pathway, the nucleotide sugars and its membrane transporters, 3'-phosphoadenosine-5'-phosphosulfate (PAPS) and Golgi trafficking proteins. Additionally, the binding specificities of the biosynthetic enzymes towards the sequences that are formed in HS growing chains influence polymerization and modification steps (133). Hence, the structural features resulting from each stage of HS biosynthesis determine subsequent modification reactions, contributing to the “step-by-step” fine tuning of HS overall structure. However, not all of the HS substrates are equally modified, thus giving rise to a great HS structural variability.

An HS biosynthesis model has been proposed, in which these enzymes are in tight interaction with each other, forming a supramolecular complex or multiple complexes termed “GAGosome”. These complexes ensure quick and concerted reactions towards the formation of GAG chains (134). However, the regulatory mechanisms underlying the activity of these possible enzymatic machinery complexes and their impact in GAGs structural variability are not yet disclosed.

As previously mentioned, HS biosynthesis occurs at the ER-Golgi interface and in the Golgi apparatus. Most HS biosynthetic enzymes are type II transmembrane proteins that are anchored to the Golgi membrane and perform their catalytic tasks towards the targeted substrates in the lumen of this organelle (135). The Golgi stack can be divided in three main compartments, including the cis-, medial- and trans-cisternae. HS enzymes localize at the center of the Golgi units, mainly cis and medial compartments, but they are not all present within the same zone of the Golgi (90). In addition, these enzymes can rapidly change their spatial distribution influencing substrate modifications (136), whereas vesicular trafficking

allows ER-Golgi cycling and re-distribution of the different enzymes (137). Overall, in addition to the expression of HS biosynthetic enzymes, their temporal and spatial distribution in the ER-Golgi might contribute to modulate HS final structures.

1.5.1. Tetrasaccharide linker formation

HS biosynthesis starts with the formation of a universal tetrasaccharide linker covalently attached to the core protein of all HS and CS/DS bearing PGs (Fig. 9). The assembly of this tetrasaccharide is initiated by two *O*-Xylosyltransferases (XYLT1 and XYLT2) transferring a Xyl unit to the Ser residues on Ser-Gly consensus regions within PGs protein moiety (138). A Gal monosaccharide is then subsequently added by the Galactosyltransferase-I/β4-Galactosyltransferase 7 (β4GALT7) (139), followed by the transient phosphorylation of the Xyl residue mediated by the kinase FAM20B. This modification is crucial for the next assembly reaction, as it enhances the activity of the Galactosyltransferase-II/β3-Galactosyltransferase 6 (β3GALT6), which will then add the second Gal residue to the nascent polysaccharide chain (140-142). The tetrasaccharide linker formation is then completed with the addition of a GlcA by the Glucuronyltransferase-I (β3GAT3), to form the linker GlcAβ1-3Gal-β1-3Gal-β1-4Xyl-β1-*O*-Ser (143). This last step takes place concomitantly with the Xyl residue dephosphorylation by the 2-Phosphoxylose phosphatase (XYLP), which is revealed to be an important step for the following polymerization of HS chains (144).

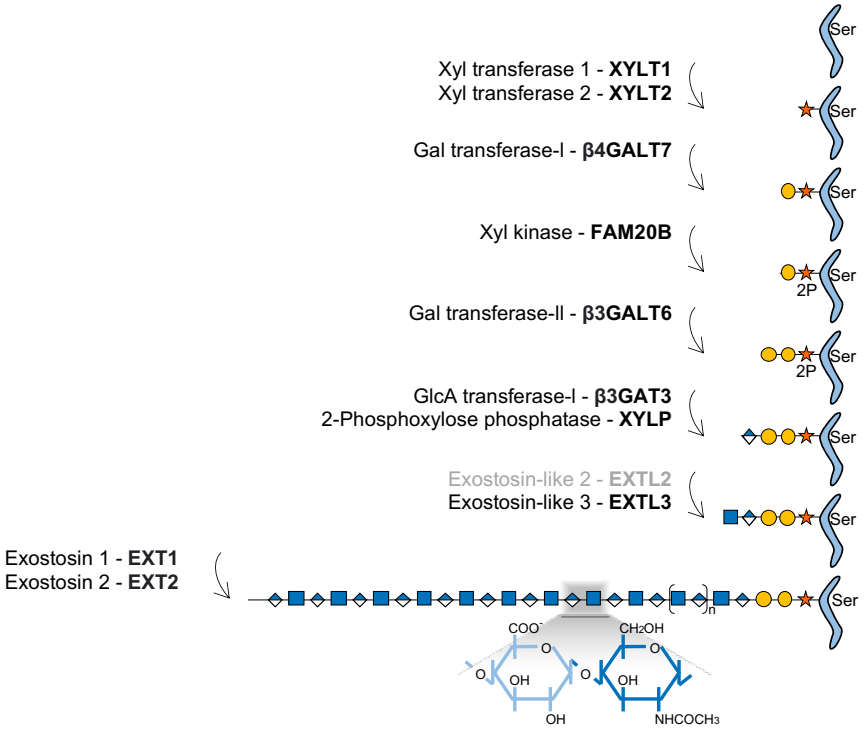


Figure 9. Schematic representation of the Heparan Sulfate (HS) biosynthetic pathway initial steps, namely tetrasaccharide linker formation and HS initiation and polymerization. The structural features of non-modified HS disaccharides were also depicted bellow the illustration. Glycan structures represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

During the synthesis of this GAG-linker, in addition to the phosphorylation of the Xyl residue, other modification reactions take place, including the sulfation of the two Gal residues (145, 146). Although the knowledge on the role of these modifications is still limited, it is speculated that these are important for the regulation of the linker biosynthesis and that they might influence the nature of the growing GAG chain (HS *vs* CS/DS). The fact that these modifications are have been mostly associated to the synthesis of CS/DS chains further favors this hypothesis (147).

1.5.2. HS initiation and elongation

At this next stage, the synthesis of a particular GAG chain is dictated by the following reaction events. The specific synthesis of HS involves the addition of the first GlcNAc residue to the mature tetrasaccharide linker, an event regulated by the Exostosin-Like 2 (EXTL2) (148) and EXTL3 glycosyltransferases (149), and is followed by further chain elongation promoted by a hetero-oligomeric complex formed by Exostosin 1 (EXT1) and EXT2, which catalyzes the alternate transfer of GlcNAc and GlcA residues (Fig. 9) (150). Noteworthy, the EXTL family includes a third homologue protein, named EXTL1, yet the role it plays in HS biosynthesis still remains poorly understood.

While EXTL3 has been thoroughly described as an efficient α1,4-GlcNAc transferase targeting the mature linker structure (149), considerable doubts were raised regarding the catalytic activity and role of EXTL2 over the years. This glycosyltransferase displays dual α1,4-*N*-acetylhexosaminyltransferase *in vitro* catalytic activity, by adding both GlcNAc and GalNAc residues to linker mimetics. Nevertheless, Kitagawa *et al*. have demonstrated that it cannot transfer GlcNAc residues to mature tetrasaccharide linker substrates (151), and that it targets instead phosphorylated forms of this linker (GlcAβ1–3Galβ1–3Galβ1-4Xyl(2-*O*-phosphate)-β1-*O*-Ser), to which adds a GlcNAc unit. By performing this catalytic step, EXTL2 promotes the synthesis of phosphorylated pentasaccharides (GlcNAcα1-4GlcUAβ1–3Galβ1–3Galβ1-4Xyl(2-*O*-phosphate)-β1-*O*-Ser) that neither EXT1 nor EXT2 can further polymerize, leading to premature HS chains termination (148).

Regarding the polymerizing enzymes, *in vitro* studies revealed that EXT1 and EXT2 display GlcA and GlcNAc transferase activities, which are considerably augmented when exerted by the hetero-oligomeric complex EXT1/EXT2 (152, 153). Interestingly, McCormick *et al.* revealed that the formation of this complex is not only important for enzymatic activity, but it might also account for the trafficking of these enzymes from ER and retention in the cis-Golgi complex, where they were shown to colocalize (150).

1.5.3. HS modification

Polymerized HS chains then undergo extensive processing and modification reactions that give rise to highly diverse fully mature HS chains (Fig. 10).

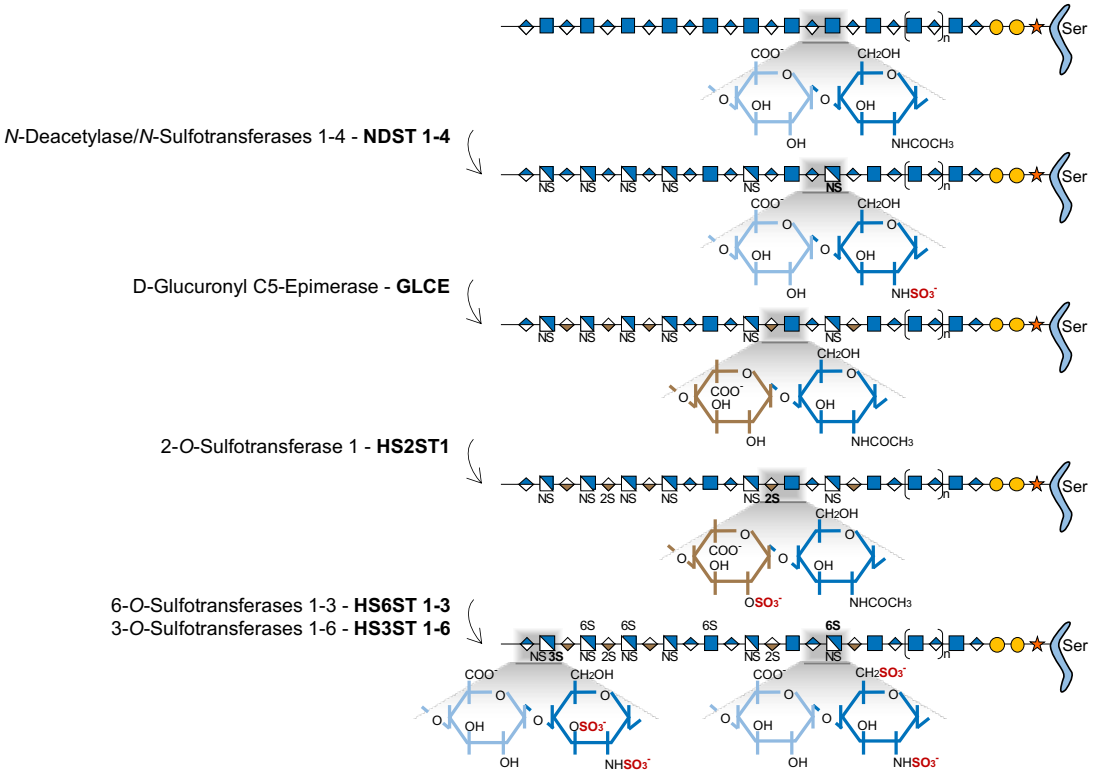


Figure 10. Schematic representation of the modification reactions occurring during Heparan Sulfate (HS) biosynthesis. The structural features of distinct HS disaccharides resulting from each modification reaction were also depicted bellow the illustration. Glycan structures represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

1.5.3.1. *N*-Deacetylation and *N*-Sulfation

During the HS modification stage, native HS polysaccharide chains, consisting of unmodified repeating GlcNAc residues linked to GlcA, are initially targeted by *N*-Deacetylase/*N*-Sulfotransferases 1-4 (NDSTs 1-4) that remove *N*-acetyl groups from GlcNAc residues and transfer a sulfate group to the generated free amino groups to form *N*-sulfated glucosamine residues (GlcNS) (Fig. 10). This first modification is essential to all the remaining reactions that take place in the Golgi, since most of the modifying enzymes act on motifs containing GlcNS (134). Of note, a small subset of *N*-deacetylated residues do not undergo *N*-sulfation, giving rise to *N*-unsubstituted glucosamine molecules ($GlcNH_3$), which elicit specific biological functions (4, 133). In vertebrates, NDST1 and NDST2 are the most widely expressed isoforms, occurring in several tissues, while NDST3 and NDST4 are more expressed at the embryonic stage and adult brain (154). The different isoforms also vary in their activity, NDST1 and NDST2 have similar *N*-deacetylase/*N*-sulfotransferase ratio activity, while NDST3 and NDST4 exhibit opposite trends, with NDST3 presenting the

lowest sulfotransferase capacity (154) and NDST4 the lowest *N*-deacetylase activity (155) of all four enzymes.

1.5.3.2. Epimerization and 2-*O*-Sulfation

HS chains are subsequently modified by D-Glucuronyl C5-Epimerase (GLCE), which catalyzes the epimerization of D-GlcA residues into L-IdoA (Fig. 10) (156). This enzyme targets specifically GlcA residues located at the reducing side of GlcNS residues and catalyzes both irreversible and reversible epimerization, since it can also convert IdoA units back to GlcA (157). This "two-way" catalytic activity was observed through *in vitro* enzymatic assays and it was shown to depend on HS *N*-sulfation patterns (GlcNS *versus* GlcNAc content) (158).

The 2-*O*-Sulfotransferase 1 (HS2ST1) catalyzes then the transfer of sulfate groups to the C2-position of both GlcA and IdoA residues, although with significantly increased efficiency towards the later, displaying around five-fold higher affinity to epimerized IdoA units (Fig. 10) (159). 2-*O*-sulfation of IdoA residues prevents their reverse epimerization into GlcA, since GLCE cannot use IdoA2S as substrate, which promotes accumulation of these residues on HS chains (160). In addition, it has been proposed that 2-*O*-sulfated GlcA residues (GlcA2S), generated at a smaller extent in HS chains, might result from early 2-*O*-sulfation, prior to the epimerization of GlcA, blocking the subsequent activity of GLCE and conversion of those residues into IdoA (161). Moreover, as a prerequisite for HS2ST1 efficient activity, the targeted GlcA or IdoA residues must also be flanked by GlcNS residues (162).

These structural alterations that ultimately dictate HS GlcA/IdoA ratio are of extreme relevance for HS functional diversity, since GlcA to IdoA epimerization confers higher structural flexibility to HS chains, which is reflected on HS-protein binding properties (156).

1.5.3.3. 6-*O*-Sulfation and 3-*O*-Sulfation

In the later stages of HS biosynthesis, 6-*O*-Sulfotransferases 1-3 (HS6STs 1-3) and 3-*O*-Sulfotransferases 1-6 (HS3STs 1-6) add sulfate groups to the C6- and C3-positions of glucosamine (GlcN) residues, respectively, contributing to the formation of more heterogeneous HS structures (Fig. 10) (163). HS6STs have broader substrate recognition and 6-*O*-sulfation activity. Enzymatic assays performed with recombinant mouse HS6STs 1-3 have shown that all isoforms can transfer sulfate groups to both GlcNAc and GlcNS residues attached to either GlcA or IdoA residues, although they target preferably IdoA-containing disaccharides (164). Furthermore, it was demonstrated that 2-*O*-sulfation of IdoA residues did not affect HS6ST activity (164). Interestingly, a different study using heparin-based oligosaccharide libraries to assess HS6ST preferred substrates, revealed that both

HS6ST2 and HS6ST3 had higher specificity for oligosaccharides with higher content in 2-*O*-sulfation (165). Habuchi H. *et al.* have also reported that different HS6ST isoforms expressed in mice had different specificities towards varied HS polysaccharide samples depending on their hexuronic acid content: HS6ST1 targeted preferentially substrates containing IdoA-GlcNS disaccharide units, HS6ST2 could act on both IdoA-GlcNS or GlcA-GlcNS motifs depending on substrate concentration, and HS6ST3 utilized both substrates independently of substrate abundance (166). *HS6STs* differential expression over different mouse organs was also reported, which supports tissue-dependent expression and consequent varying effects on HS structural motifs (166). 6-*O*-sulfation contributes to fine-tune the sequential biosynthesis of HS chains, since GlcA epimerization by GLCE is precluded by 6-*O*-sulfation of adjacent GlcN residues (160), and 2-*O*-sulfation of IdoA by HS2ST1 is inhibited if the adjacent GlcN residue is 6-*O*-sulfated (162, 167). These data are in line with the described stepwise mechanism underlying HS biosynthesis, further supporting that epimerization and 2-*O*-sulfation reactions usually take place prior to 6-*O*-sulfation. More recently, it was uncovered that 6-*O*-sulfated HS oligosaccharides were able to bind to HS2ST1, with higher affinity than HS2ST1 direct substrates, and inhibit its catalytic activity (168). In light of these results, it was hypothesized that enzymes involved in the HS biosynthetic pathway might be temporarily and/or spatially separated from the intermediate substrates produced downstream on the Golgi. This would explain why HS modified structures do not impair the activity of initially acting sulfotransferases towards new native glycans (168).

Although HS3STs constitute the largest family of HS sulfotransferases, composed by 7 isoforms (HS3ST-1, -2, -3A, -3B, -4, -5 and -6), GlcN 3-*O*-sulfation is the least frequent modification (90). The different isoforms exhibit specific binding affinity to GlcN residues linked to different hexuronic acid residues contributing to the functional diversity of HS chains. HS3ST1 transfers sulfate groups to GlcNS and GlcNS6S bound to unsulfated GlcA and IdoA units (169, 170). *In vitro* enzymatic assays have revealed that murine HS3ST1 activity is specifically inhibited by IdoA2S residues linked to the targeted GlcN residues, suggesting an additional layer of regulation, according to which HS 2-*O*-sulfation content impacts 3-*O*-sulfation patterns and overall levels (170). HS3ST2 transfers sulfate groups to GlcNS residues attached to either GlcA2S or IdoA2S, while HS3ST3A and HS3ST3B isoforms act on GlcNS linked to IdoA2S (171). Later, it was also shown that the HS3ST3A isoform can utilize substrates on rarer regions of HS motifs, including $GlcNH_3$ and N-unsubstituted 6-*O*-sulfated glucosamine ($GlcNH_26S$) residues linked to IdoA2S at the non-reducing end (172). HS3ST4 and HS3ST6 catalyze 3-*O*-sulfation on similar disaccharides targeted by HS3ST3 isozymes (GlcN residues linked to IdoA2S residues at the non-reducing site) (173, 174). Lastly, HS3ST5 was revealed to have a broader substrate

specificity when compared with the other isoforms, targeting both $GlcNH_3$ and GlcNS residues at the reducing end of GlcA, IdoA and IdoA2S residues (175).

1.5.4. HS Post-Synthetic Modification

Once HS chains are fully synthesized, and mature HSPGs are expressed at the cell surface, additional post-synthesis reactions can take place and further modify HS chains, namely 6-*O*-desulfation catalyzed by 6-*O*-endosulfatase 1 (HSULF1) and HSULF2, and cleavage of the chains by Heparanase 1 (HPSE), thus generating additional structural diversity with biological relevance (Fig. 11) (90, 176).

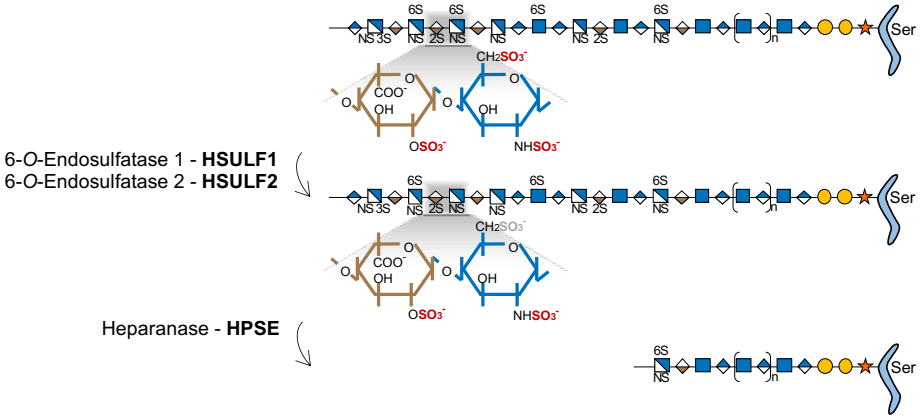


Figure 11. Schematic representation of the post-synthesis modification reactions occurring after Heparan Sulfate (HS) biosynthesis. The structural features of distinct HS disaccharides resulting from each modification reaction were also depicted bellow the illustration. Glycan structures represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

1.5.4.1. HS Cleavage

HPSE is an endo-β-D-glucuronidase that targets internal GlcUAβ1-4GlcNS linkages, cleaving HS chains in shorter fragments of 10-20 units (5-10 kDa). The presence of specific sulfation features, including *N*-sulfation and 6-*O*-sulfation, favors HS recognition by this enzyme (177, 178). HPSE is initially secreted to the extracellular environment in an inactive form – pro-HPSE – and it interacts with different cell membrane molecules, including HSPGs, to form a complex that is later endocytosed and trafficked to lysosomes. In lysosomes, HPSE is cleaved by the cathepsin L protease, resulting in the formation of a mature active HPSE that can then be secreted again (179). Besides targeting cell surface HS, HPSE can exert its catalytic functions in the ECM, contributing to ECM remodeling, the release of HS fragments, and the consequent discharge of HS-pooled soluble ligands (180).

An HPSE homolog enzyme termed HPSE2 has been identified and characterized later, and although it shares ~40% homology with HPSE and is able to bind HS chains with high affinity, it lacks most of the glucuronidase catalytic domain, being unable to cleave HS (181, 182). Furthermore, HPSE2 inhibits HPSE activity, possibly *via* HS competitive binding, and is proposed to serve as a HPSE antagonist in multiple pathologies (183, 184) .

1.5.4.2. HS 6-*O*-Desulfation

HSULF1 and HSULF2 are described as important 6-*O*-endosulfatases that are responsible for shaping HS 6-*O*-sulfation motifs extracellularly, after biosynthesis. Both HSULFs are unique human sulfatases in the sense that they catalyze the hydrolysis of 6-*O*-sulfate groups in the C6- position of GlcNS6S residues located in the internal regions of the chains. They target preferentially IdoA/GlcA2S-GlcNS6S trisulfated disaccharides, though low catalytic activity against IdoA/GlcA-GlcNS6S disulfated disaccharides has also been reported (185). HSULFs act through an oriented and processive mechanism, targeting the non-reducing end of the highly sulfated S-domains and acting towards its reducing end (186). These enzymes are secreted to the extracellular environment in a pre-protein format and are then subjected to proteolytic cleavage by furin-type proteases to achieve a mature active state. Despite presenting limited activity, catalyzing the removal of few 6-*O*-sulfate groups within highly sulfated regions of HS, HSULFs have great biological consequences (187).

The expression of HSULFs has been differentially detected in several human tissues, as *HSULF1* appeared to be highly expressed in testes, stomach, skeletal muscle, lung, and kidney, whereas higher levels of *HSULF2* were reported in ovary, skeletal muscle, stomach, brain, uterus, heart, kidney, and placenta (188) Noteworthy, Seffouh *et al.* showed that despite both isoforms displaying a similar catalytic mechanism for 6-*O*-desulfation, HSULF1 acted at a higher rate when compared to HSULF2 (186). In addition, *in vivo* assays seem to suggest that both isoforms might present redundant functions, as double *Sulf-1/Sulf-2* KO mice present high neonatal mortality and serious morphological defects, while either *Sulf-1* or *Sulf-2* KO mice did not present such severe phenotypes during development (189). Interestingly, there is evidence of an interplay between biosynthetic enzymatic activity and post-synthesis modification mechanisms, as Lamanna *et al.* showed that *Sulf-1* and/or *Sulf-2* KO mice displayed HS structural changes that were not dependent directly on the Sulfs enzymatic activity, such as altered *N*- and 2-*O*-sulfation (190).

1.5.5. CS initiation, elongation and modification

As previously alluded, GAGosylation of all HS and CS/DS bearing PGs is initiated with the formation of a tetrasaccharide linker, which primes for either HS or CS polymerization. Currently it remains to be fully elucidated which are the molecular mechanisms determining the synthesis of a specific GAG chain in detriment of the other. In addition to the linker modifications that might contribute to fine-tune this process, it has been proposed the existence of specific amino acid sequences in the vicinity of proteins' Ser-Gly sequons that might have an important role in HS/CS assortment. The presence of two or more nearby Ser-Gly dipeptides, close to a group of acidic amino-acids and a tryptophan residue was

reported to favor HS synthesis (191, 192), although more distant PG protein elements have also been reported to be important determinants for preferential HS formation (193). Recent work by Sammon *et al.* might have further lifted the veil on this intricate process. *In vitro* analysis indicated that a biosynthetic enzyme that initiates CS biosynthesis (CS *N*-acetyl-galactosaminyltransferase 2; CSGALNACT2) tends to recognize the linker, rather than the modified polypeptide, and act in a more promiscuous manner, modifying both HS and CS PGs, whereas the HS glycosyltransferase EXTL3 interacts with the linker and the polypeptide backbone, only targeting specific protein entities that are indeed present in nature as HSPGs (45). Bourgeais *et al.* have further validated EXTL3 as pivotal element at the center of this decision (194). These results further highlight the importance of biosynthetic enzymes distribution within the Golgi.

Briefly, the polymerization of CS chains is initiated by CS *N*-acetyl-galactosaminyltransferase 1 (CSGALNACT1) and CSGALNACT2, adding the first GalNAc residue to the linker (195). Polymerization of CS is then continued by four polymerizing enzymes, namely CS synthase 1 (CHSY1), CS synthase 3 (CHSY3), CS synthase 2 (CHPF) and CS glucuronyltransferase (CHPF2), that will transfer repeating GalNAc and GlcA residues in an alternate manner (Fig. 12). These four enzymes can form multiple enzyme complexes to promote CS elongation, each complex interacting differently with the protein backbone, leading to the formation of CS chains of variable length in a process that remains to be fully disclosed (196-198).

Following or during CS elongation, similarly to what happens during HS maturation, different sulfotransferase enzymes catalyze the *O*-sulfation of CS disaccharides at multiple positions (Fig. 12). The Chondroitin 4-*O*-sulfotransferases (CHST11-13) and Chondroitin 6-*O*-sulfotransferases (CHST3 and 7) catalyze the transfer of sulfate groups to the C4- and C6-position of GalNAc residues to form CS A (GlcA-GalNac(4S)) and CS C (GlcA-GalNac(6S)) units, respectively. The GalNAc 4-*O*-sulfate 6-*O*-sulfotransferase (CHST15) targets this CS A units and transfers sulfate groups to the GalNAc C6-position, forming CS E (GlcA-GalNac(4S6S)). The uronyl 2-*O*-sulfotransferase (UST) then catalyzes GlcA 2-*O*-sulfation, forming CS D (GlcA(2S)-GalNac(6S)) units.

Additionally, epimerization of GlcA into IdoA residues catalyzed by the DS-epimerase 1 (DSE) and 2 (DSEL) might occur in non-sulfated disaccharides, leading to the formation of DS chains (Fig. 12). These GAGs can also be modified by the Dermatan 4-*O*-sulfotransferase (D4ST/CHST14), the CHST15 and the UST to form DS iA (IdoA-GalNAc(4S)), iE (IdoA-GalNAc(4S6S)) and iB (IdoA(2S)- GalNAc(4S)) units (199, 200).

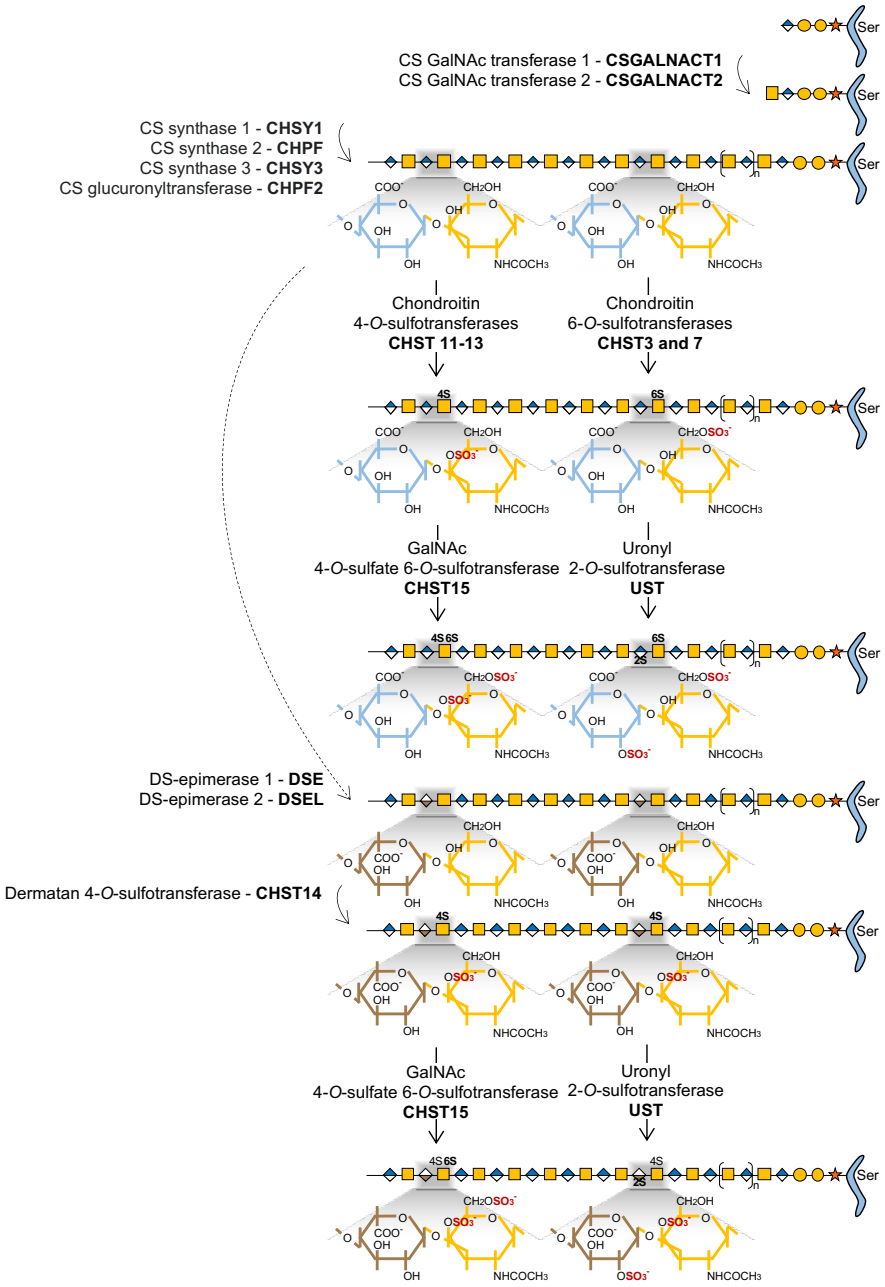


Figure 12. Schematic representation of CS/DS biosynthesis, from the initial steps of CS polymerization to CS/DS modification. The structural features of distinct CS disaccharides resulting from each reaction were also depicted bellow the illustration. Glycan structures represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

1.6. Heparan Sulfate biological roles and ligand binding specificities

Overall, the full set of reactions catalyzed by HS biosynthetic and editing enzymes determine glycan chain length, epimerization and sulfation profiles, resulting in the synthesis of HS chains with extremely high structural variability (90). Regardless, all HS chains commonly feature two main types of structural domains: sulfated (S)-domains, enriched in highly modified disaccharides, i.e. sulfated hexuronic acid units linked to O-sulfated GlcNS residues, that are successively intercalated by *N*-acetylated (NA)-domains

bearing mostly non-modified GlcNAc units linked predominantly to GlcA residues (Fig. 13) (201, 202).

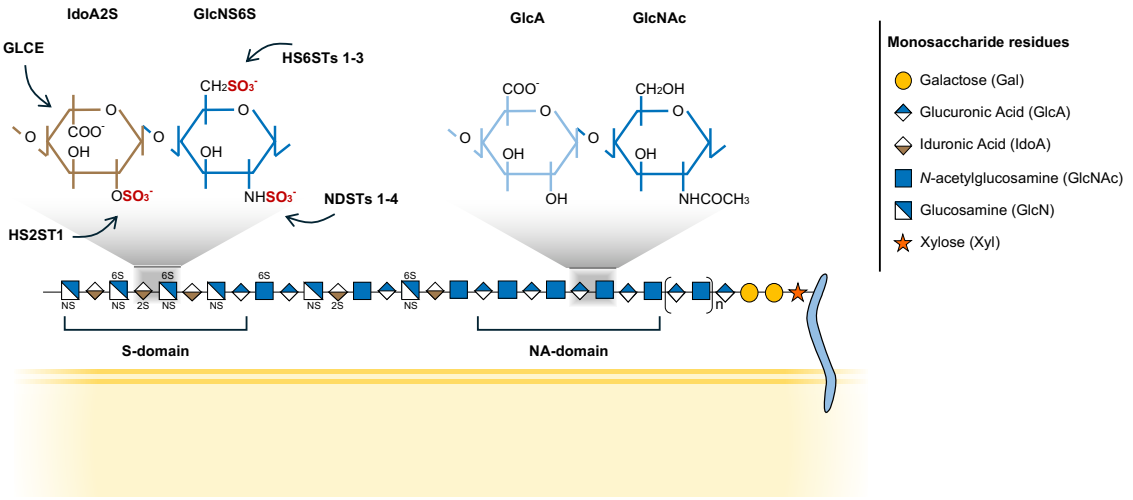


Figure 13. Illustrative representation of a mature Heparan Sulfate (HS) chain and its structural organization. In mature HS chains, HS disaccharides are organized in two main structural domains: sulfated (S)-domains, enriched in highly modified disaccharides, and *N*-acetylated (NA)-domains, that are mostly composed by non-modified GlcNAc units linked predominantly to GlcA residues. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) (1).

Most biomolecules interact with HS S-domains, rendering these domains responsible for most of the polysaccharide biological activities. The high negative charge within these regions favors the non-covalent ionic bonding between HS and positively charged amino acid residues in multiple protein targets. These features indeed modulate HS binding specificities to protein targets, and vary over different organs (203), developmental stages (204-206) and pathologies (207-209), allowing HSPGs to participate in many signaling cascades and take over a range of cell regulatory events.

HS chains bind to multiple biologically active molecules, including growth factors, chemokines, cytokines, morphogens, extracellular structural proteins, enzymes involved in different biochemical pathways and transmembrane signaling receptors (87, 210). This prompts cell surface HSPGs to modulate main cellular events, by acting as co-receptors or scaffolds for protein-protein interactions, triggering receptors activation and subsequent signaling transduction, and as important mechano-signaling transducers of extracellular stimuli (38, 94). HSPGs not only enhance the activity of neighboring receptors expressed on the same cell, but also participate in the transactivation of receptors, including key RTKs, in adjacent cells, mediating cell-cell crosstalk (211). Recently, HSPGs have been shown to play a regulatory role in the activation of calcium channels, and it has been proposed a mechanism dependent on the cytosolic calcium levels, through which HSPGs modulate cells' cytoskeleton organization and adhesion (212). In the ECM, HSPGs can also function

as storage units and contribute to the formation of extracellular gradients of varied soluble molecules, thereby modulating their availability (88).

The impact of HS overall sulfation degree versus specific sulfation patterns as determinants for HS-protein binding has been debated over a long time. It has been established that in the majority of cases tissue-dependent HS general sulfation, and sulfation distribution by blocks (S-domains and NA-domains), rather than unique sulfation sequences encoded in HS glycan chains, determine binding affinities (213). The current consensus is that *N*- and 2-*O*-sulfated units are involved in low specificity binding, 6-*O*-sulfated units in intermediate specificity binding and 3-*O*-sulfated and *N*-unsubstituted units are responsible for high specificity binding, which justifies the strong binding overlap of protein ligands towards HS with variable sulfation patterns reported in different HS-binding protein assays.

Although unique and distinctive biologically active HS structural motifs have not been disclosed for most protein-HS interactions, distinct roles of the most common types of sulfation (*N*-, 2-*O*- and 6-*O*-sulfation) have been reported for protein binding affinities and biological activities. FGF2 was the first growth factor reported to depend upon cell surface HS as a co-receptor required for the formation of a biologically active FGFR-FGF2-HS ternary complex (214, 215). FGFs constitute a large family of mitogens involved in the regulation of cellular migration, proliferation, differentiation, and survival (216). FGF2-HS interaction, in particular, is important not only for receptor dimerization and activation, but also for evasion of degradation by FGF2 molecules (217). Maccarana *et al.* showed that IdoA2S and GlcNS were important units for FGF2 binding, whereas the presence of GlcN6S residues was irrelevant for HS-FGF2 interaction (218). It was later reported that 6-O-sulfated residues were required to bridge FGF2 and FGFR, promoting the formation of the full ternary complexes and receptor activation (219-221).

Interestingly, a specific trisaccharide that included both IdoA2S and GlcNS6S residues (IdoA2S–GlcNS6S–IdoA2S) was invariably detected in FGF1 binding sites in HS chains (222). It was demonstrated that the presence of this specific motif repeated in HS polysaccharides contributed to higher FGF1 binding affinity, when compared with HS polysaccharides with increased overall sulfation, and that 6-*O*-sulfation was important for FGF1-HS interaction (222). Overall, HS structural requisites for binding of FGF1 and FGF2 illustrate the consensus on the general role of *N*- and 2-*O*-sulfation in ligand binding and the specificity provided by 6-*O*-sulfation.

HSPGs also provide a scaffold for stabilizing extracellular morphogen gradients, crucial during early development. Noteworthy, signaling pathways of major HS-binding morphogens including Shh, BMP and Wnt, were shown to be impaired by aberrant HS biosynthesis, resulting from the loss of EXT gene expression and function (223, 224). HS

bioactive domains for Wnt and for a member of the BMP family, BMP-2, have also been previously disclosed. Wnt interaction with HS and subsequent activation were associated mainly with the polysaccharide 6-*O*-sulfation content, and its structural remodeling by the HSULFs (225, 226). The most critical components of BMP-2 HS-binding domain were revealed to be GlcNS residues, while 6-*O*- and 2-*O*-sulfation were also important, but to a lesser extent (227).

The most selective interactions take place in regions of HS chains that contain rare modifications, such as $GlcNH_3$ and 3-*O*-sulfated GlcN residues (228). The first identified and most studied example of a HS specific saccharide sequence required for protein high specificity binding and activity relates to antithrombin, an important inhibitor of blood coagulation, whose inhibitory activity is potentiated upon binding to heparin (229). The heparin pentasaccharide GlcNAc6S-GlcA-GlcNS3S6S-IdoA2S-GlcNS6S has been disclosed as a specific binding epitope within heparin responsible for antithrombin interactions (230, 231), and the GlcNS3S residue within this sequence is highlighted as a fundamental distinctive structural feature as it is an otherwise rare unit within heparin/HS chains (230, 232).

This rare HS modification has also been associated with a few highly selective interactions with other types of bioactive molecules with impact in cell physiology, including the growth factor FGF7 (233), cyclophilin B (234) and the cell surface receptor Neuropilin-1 (235).

1.7. Molecular functions of Heparan Sulfate and Heparan Sulfate Proteoglycans in disease

Over the years, as the contribution of HS and HSPGs to multiple normal physiological events during development and homeostasis has become clearer, increasing awareness has been raised towards the leading roles these molecules assume in many pathophysiological processes, including infection and inflammatory diseases, neuronal and neurodegenerative disorders, genetic conditions and cancer (236).

1.7.1. Cell-surface HSPG aberrant expression in cancer

HSPGs are key players in cancer development and progression, acting as important maestros of cancer cell interactions with the ECM and cancer cell communication, ultimately controlling the tumor microenvironment biochemical and biophysical features (5, 237, 238). The expression of these macromolecules is altered in several types of cancer, contributing to the deregulation of different cell events, including proliferation, angiogenesis, adhesion, migration and invasion. Interestingly it appears that the same HSPGs can elicit contradictory effects and either promote or inhibit tumorigeneses, possibly by participating in different

signaling cascades in a cell and/or tissue dependent manner. This might explain the variations in HSPGs expression and influence patients' overall survival and prognosis in different malignant pathologies that will be described in this section, which is focused primarily on the main cell surface HSPGs.

SDC1 altered expression has been reported in different types of tumors. It was found to be upregulated in breast (239, 240) and cervical cancers (241), showing up as an indicator of poor prognosis. *SDC1* overexpression was also described in pancreatic carcinoma (242). Interestingly, Juuti *et al.* alluded to the possibility that in pancreatic cancer this HSPG might act differently depending on its location, and attributed pro-tumorigenic properties specifically to stromal SDC1 (243). Also in cervical cancer, soluble and membrane bound SDC1 had different effects over cell invasion (241).

Within the breast cancer context, induced knock-down of epithelial SDC1 inhibited FGF2 mediated mitogen-activated protein kinase (MAPK) signaling, whereas overexpression of this molecule promoted cell proliferation, suggesting this HSPG acts as a co-receptor in tumor cells pro-mitogenic phenotype (244). SDC1 has also been implied in the regulation of angiogenesis (245), inflammation and cancer stem cell-like phenotypes (246) in highly aggressive forms of breast cancer, namely triple negative and inflammatory breast cancer.

On the other hand, *SDC1* reduced expression in many other types of cancer has been associated with poor outcomes, including hepatocellular carcinoma (247), head and neck (248) and colorectal cancer (249).

SDC2 overexpression is a frequent feature of several tumors, namely colorectal (250), lung (251), pancreatic (252) and breast (253-255) cancer. The pro-oncogenic role of SDC2 was further validated using cell lines from these different malignancies in both *in vitro* and *in vivo* assays, showing that SDC2 promotes aggressive phenotypes by inducing cell motility, invasion, tumor growth and metastasis. These phenotypes were, in some models, associated with K-ras/MAPK (pancreatic cancer), RhoGTPases (breast cancer), or transforming growth factor-β (TGF-β) signaling (breast cancer).

In colorectal cancer, *SDC2* upregulation was associated with stage, vascular invasion, lymph node and distant metastases, and was inversely correlated with patients' overall survival. Aligned with this data, the same study showed that SDC2 knock-down in colorectal cancer cell lines inhibited cell proliferation, migration and invasion, and it inhibited the expression of epithelial-to-mesenchymal transition (EMT) markers, suggesting that it plays a role in the initial stages of cell transformation and tumor progression (250).

SDC3 in the context of cancer has been significantly less explored. Fewer reports have identified upregulation of this molecule in pancreatic cancer, specifically localized in neuritis and perineurium of pancreatic nerves (256), and in ovarian cancer (257). Moreover, *SDC3*

high expression has also been correlated with better prognosis of melanoma patients, but particularly associated with tumors exhibiting high hypoxia profiles (258).

SDC4, as an important transmembrane HSPG that participates in a wide variety of signaling processes, from facilitating ECM-cytoplasm cross-talk and modulating intracellular signaling, to acting as a scaffold protein, has been identified as key participant in multiple cell events related to oncogenesis and tumor progression (259).

In lung cancer, *SDC4* upregulation induces cell migration and invasion, along with EMT, mainly by modulating TGF-β1 cascade (260). Noteworthy, it was described more recently that *SDC4*, which is upregulated in high-grade osteosarcoma (261), is also responsible for maintaining TGF-β1 signaling in osteosarcoma lung metastasis, similarly contributing to EMT and cell invasion (262). In hepatocellular carcinoma, SDC4, together with SDC1, were determined to be essential for chemokine-mediated cell motility and invasion (263).

SDC4 upregulation was found in testicular cancer (264) and invading colorectal tumor cells, where it associated with patients' poor overall survival (265). In addition, it is an important biomarker in glioblastoma, with predictive value for the selection of patients responsive to immunotherapy (266). In breast cancer, there is contradictory information on the expression and impact of this molecule. Baba *et al.* described that *SDC4* expression positively correlated with grade, and size of primary tumors (267), whereas in a later study *SDC4* reduced expression was associated with high grade breast tumor (268). More recently, Onyeisi *et al.* reported that high SDC4 expression was a positive factor for patient survival. However, in a subset of estrogen receptor negative and estrogen/progesterone-receptor negative patients it associated with poor survival (269). The authors further described a new regulatory pathway for *SDC4* expression, controlled by the microRNA-140-3p, and showed that *SDC4* silencing reduced breast tumor cells' invasion and migration capabilities, and enhanced *in vitro* cell-matrix adhesion (269).

SDC4 expression has also been determined to negatively impact gastric carcinogenesis and to contribute to disease progression and patients' poor prognosis (270). This will be further detailed in the section 1.7.3. of this chapter.

The members of the GPC family have also been frequently implicated in several tumorigenesis events. *GPC1* high expression has been reported in pancreatic cancer and associated to cell proliferation and metastasis, angiogenesis, and patients' overall poor prognosis (271-273). In addition, a study has reported that GPC1 positive EVs could be detected in blood samples of pancreatic cancer patients and possibly serve as highly specific biomarkers (274). Likewise, *GPC1* is overexpressed in breast cancer (275) and glioblastoma (276). In breast cancer, GPC1 was shown to be crucial for heparin-binding EGF-like growth factor (HBEGF) and FGF2 mediated cell proliferation.

GPC2 is mostly highly expressed in neuroblastoma, correlating with patients' poor prognosis. Also in this work, it was demonstrated that GPC2 drives neuroblastoma tumor cell growth *via* the regulation of the Wnt/β-catenin signaling pathway (277). Similarly, overexpression of *GPC3* has been reported in hepatocellular carcinoma (278), where it also plays a role as Wnt co-receptor, although in a GAG independent manner, promoting disease progression by stimulating the canonical Wnt signaling (279).

There is considerably less data on the expression of *GPC 4-6* in cancer pathologies. *GPC4*, like *GPC1*, has also been reported as overexpressed in pancreatic cancer, and associated with patients' poor prognosis (280), *GPC5* is overexpressed in rhabdomyosarcoma (281) and *GPC6* upregulation has also been detected in ovarian cancer (282). Finally, and in contrast with the general trend towards GPCs expression in cancer, *GPC5* was revealed to be downregulated in prostate cancer (283).

1.7.2. Heparan Sulfate expression levels and sulfation features in cancer

Altered cellular levels of HS, not only resulting from aberrant HSPGs expression but also from the abnormal expression of enzymes involved in HS biosynthesis and editing, are main features heavily involved in cells' malignant transformation (5, 94).

1.7.2.1. HS glycosyltransferases deregulation

Regarding HS biosynthetic enzymes, Karibe *et al.* reported *EXTL3* promoter hypermethylation, and consequent gene silencing, accompanied by loss of HS in tissue samples from mucinous colorectal cancer patients (284). The expression of several HS glycosyltransferases, namely *EXT1*, *EXT2*, *EXTL1*, *EXTL2* and *EXTL3*, has also been screened in a set of estrogen receptor positive and estrogen/progesterone receptor and HER2 triple negative breast cancer cell lines. Expression of *EXT2* and *EXTL3* was reduced in all tested cell lines, while *EXTL2* was upregulated in triple-negative cell lines (285). Interestingly, these changes were not reflected in the levels of HS determined in the different cell lines, suggesting that other biosynthetic enzymes or altered expression of different HSPGs might overcome the role of HS glycosyltransferases in determining overall HS levels. The impact these specific changes might have in carcinogenesis remains to be explored.

1.7.2.2. Sulfotransferase expression deregulation and aberrant HS sulfation profiles

Comparative studies have demonstrated aberrant expression of several genes encoding HS biosynthesis machinery in cancer. Transcriptomic and immunohistochemical analyses have been performed in breast (209) and colorectal (286, 287) tumor samples and

indicated significant variations in the expression and tissue- distribution of several enzymes involved mostly in HS epimerization and sulfation. In fact, most variations in the expression of HS-related genes detected in several cancer pathologies concern enzymes specifically involved in HS modification reactions that change HS sulfation degree and patterns. This is likely justified by the dependence of the numerous biological roles displayed by HSPGs on these structural features, as discussed previously.

Regarding NDST isozymes that act in the initial step of HS modification, namely GlcNAc *N*-deacetylation and *N*-sulfation, loss of *NDST4* gene expression in colorectal cancer has been associated with higher pathological stages and patient's poor prognosis (288). Additionally, Fuster *et al.* studied the impact of another NDST isoform, NDST1, on tumor angiogenesis, and showed that endothelial cells isolated from *Ndst1* KO mice synthesized structurally modified HS chains that impaired angiogenesis-related signaling pathways, leading to decreased vascularization of lung tumors and consequent reduced tumor growth (289). Altered expression of *GLCE*, the epimerase controlling HS GlcA/IdoA ratio, was also reported in cancer. Studies indicated significant decreased expression of *GLCE* in breast tumors (290), even at premalignant stages of the disease, and in lung cancer cells (291). Furthermore, anti-proliferative effects were reported on both breast (292) and small-cell lung (291) cancer cells after induced re-expression of *GLCE*, highlighting this gene as a potential tumor-suppressor gene.

The sulfotransferases that catalyze the more common HS *O*-sulfation reactions, HS2ST1 and HS6STs, have also been implicated in cancer pathology. Gene expression datasets from the Oncomine database were explored and revealed that *HS2ST1* expression is upregulated in prostate carcinoma and functional assays showed that it correlates with cell proliferation and invasion capabilities, known to enhance cancer cells metastatic behavior (293). In the same line, but in a different model, it was demonstrated that downregulation of *HS2ST1*, and concomitant change in HS structural patterns, accompanied the granulocytic differentiation of SKM-1 leukemia cells and were associated with cells' growth inhibition and less aggressive phenotypes (294). In breast cancer, *HS2ST1* expression was shown to promote the acquisition of cancer stem cell-like properties, by activating stemness-associated signaling pathways (295). Additionally, and contradicting the previously reported HS2ST1 "tumor promoter gene"-like features, upregulation of *HS2ST1* in breast cancer cell lines was shown to promote cell-cell and cell-ECM adhesion and to inhibit cell migration and invasion capabilities (296). These effects were associated with increased HS 2-*O*-sulfation and subsequent altered growth factor binding affinities, decreased expression and activation of the EGFR, which impaired different signaling pathways that modulate cancer cell invasiveness, and increased expression of E-cadherin (296).

Regarding HS6STs, Waaijer *et al.* reported upregulated expression of *HS6ST1* and *HS6ST2* during chondrosarcoma progression, correlated with increasing tumor histological grade, higher levels of HS6ST3 in most of the analyzed cartilage tumors, and higher levels of 6-*O*-sulfated HS disaccharides in high-grade chondrosarcoma cell lines (297). Similarly, the isoforms *HS6ST2* and *HS6ST3* were also found to be overexpressed in colorectal (298) and breast (299) cancer, respectively. *HS6ST3* expression in particular was shown to influence breast tumor cell growth, invasion and migration (299). HS 6-*O*-sulfation content was also shown to be augmented in the ovarian endothelium of patients with advanced stage ovarian cancer (300). In agreement with this work, it was later disclosed the impact of HS 6-*O*-sulfation content, majorly determined by HS6ST isozymes, on ovarian cancer angiogenesis (301). In this work, Cole *et al.* showed that HS 6-*O*-sulfation stimulates the HBEGF-dependent activation of cell surface receptor EGFR, which leads to increased expression of angiogenic cytokines (interleukin 6, interleukin 8 and FGF2) by ovarian tumor cells (301).

Lastly, aberrant expression of HS3STs, which catalyze the rarest HS modification, was also reported in various human cancers, suggesting an important role in cell malignant transformation and tumor cells' functional features. Miyamoto *et al.* described hypermethylation of 5' region of the *HS3ST2* gene and consequent loss of its expression in human primary breast, colon, lung and pancreatic cancers (302). Later, Kumar *et al.* have taken this premise and further investigated the possible functions of HS3ST2 in tumorigenesis (303). These authors reported that re- expression of this gene in a highly invasive breast cancer cell line augmented the activation of MAPK and Wnt/b-catenin signaling pathways, which promoted breast cancer cell invasiveness, motility and chemoresistance (303). Additionally, *HS3ST2* expression in breast cancer cells was shown to be related with acquisition of a cancer stem cell-like phenotype (295). Hypermethylation and silencing of *HS3ST2* was also reported in non-small cell lung cancer, and it was associated with poor overall patient survival. Hwang *et al.* induced exogenous expression of this gene in two lung cancer cell lines and observed reduced cells' migration and invasion capabilities, as well as their lower proliferation rate, supporting that *HS3ST2* hypermethylation might promote lung tumorigenesis (304). Regarding different isoforms, *HS3ST3A* was revealed to be epigenetically repressed in breast cancer cell lines representative of distinct molecular subgroups, and it was shown to induce opposite effects, either oncogenic or tumor-suppressive, depending on tumor cell phenotypes (305). Moreover, *HS3ST3B1* expression was associated to acute myeloid leukemia progression, by inducing expression and shedding of proangiogenic factors (306), and, more recently, it was found to be upregulated on non-small cell lung cancer and to regulate EMT (307). Interestingly, the telomerase protein TRF2, whose expression is upregulated in tumor cells

and associated to suppression of immune response, was shown to regulate *HS3ST4* expression and to prevent natural killer (NK) cell recruitment and promote cancer immune escape (308, 309).

Together with changes in HS modifying enzymes, altered HS sulfation levels and structural conformations have also been studied and described in different types of cancer, in an attempt to profile specific sulfation patterns as distinctive traits in cancer pathologies that can potentially serve as predictive biomarkers.

GAG disaccharide profiling of cancer patients' tissue samples has revealed decreased HS 6-*O*-sulfation and total *O*-undersulfation in hepatocellular cancer (310) and overall decreased HS sulfation, both *N*- and *O*-, in renal cell carcinoma, mainly due to lower levels of mono 6-*O*-sulfated disaccharides, and all di-sulfated disaccharides (311). HS sulfation was also shown to vary in breast cancer but in a cell dependent manner. As such, it was shown that MCF7 breast cancer cells, expressing estrogen and progesterone receptors, produce HS with reduced content in 2-*O*-sulfated disaccharides and enriched in 6-*O*-sulfation. In contrast, HS isolated from MDA-MB-231 and HCC38 cells, triple-negative for estrogen and progesterone receptors, as well as HER2, were more 2-*O*-sulfated and less 6-*O*-sulfated, both compared with a non-tumorigenic human mammary gland epithelial cell line (285).

On the other hand, data regarding 3-*O*-sulfation levels in tumors is very scarce. This is mainly due to the absence of commercially available 3-*O*-sulfated disaccharide standards that can be used for HS disaccharide structural analysis, as well as the low susceptibility of 3-*O*-sulfated disaccharides to the currently used Heparinases (I, II and III) for exhaustive HS depolymerization, which is a prerequisite for this type of analysis.

1.7.2.3. Sulfatase activity deregulation

HS post-biosynthesis editing by 6-*O*-endosulfatases represents an additional regulatory mechanism frequently altered in different tumor models, further emphasizing the role of polysaccharide altered 6-*O*-sulfation in cancer (312, 313). HSULF1 is often described as having a tumor suppressor effect while HSULF2 exhibits generally pro-oncogenic properties, although this trend is variable depending on the studied pathology. HSULF1 was shown to be downregulated in head and neck squamous cell carcinoma cell lines (314) and hepatocellular carcinoma (315). On the other hand, it was upregulated in pancreatic cancer cells (316), leukemia and renal carcinoma (317). In the meantime, HSULF2 upregulation in cancer has been more consistently reported in lung (318), pancreatic (319), breast (320), colorectal (321), and hepatocellular cancer where its expression was associated with patients worst prognosis (322). It remains unclear how two different isoforms of the same enzyme, with overlapping catalytic activity and substrate specificities can exhibit different

expression pattens and roles in cancer. Interestingly, it was recently discovered that HSULF2, but not HSULF1, is modified with a CS/DS chain, classifying this enzymes as an extracellular, secreted part-time PG. The presence of the CS/DS chain was demonstrated to fine tune its enzymatic activity and to modulate HSULF2-mediated tumor growth and metastasis in cultured cells and *in vivo* (323). The authors have shown that the removal of this GAG chain was important to activate HSULF2 catalytic activity, and postulated that this GAG-based enzymatic regulation might contribute to the differential roles, and impact in disease outcomes, attributed to each of these enzymes.

1.7.3. Heparan Sulfate Proteoglycans and Heparan Sulfate aberrant expression in gastric cancer

Given the significant role of HS in various malignant pathologies, numerous studies have been conducted over the last 15 years to identify the expression profile and predictive value of HS in gastric cancer. However, divergences in the reported data and the heterogeneity of results have obstructed the successful completion of this challenging but fundamental task. Initially, through HPLC-based GAG structural analyses, increased CS/DS and HA levels and decreased HS content was described in gastric cancer (208). Later, Lo *et al.* have reported HS detection in both stroma and epithelia by employing immunohistochemical analyses. In addition, and aligned with the beforementioned work, decreased HS epithelial expression was associated with higher tumor grades, greater depth of tumor infiltration and reduced patient survival after recurrence (324). Finally, and employing mass spectrometry-based GAG disaccharide structural analysis, Weyers *et al.* have failed to confirm this trend, and reported no variations in HS levels and sulfation. They described instead, significant increase in non-sulfated CS/DS content and a decrease in HA levels in gastric cancer (325). The heterogeneity of patient tissue samples, the differences in the used cohort sizes and in the sensitivity of the used methodologies might explain some of the observed differences. Moreover, evaluating more carefully the distribution of both HS and HS carriers in tissue samples and the combination of this data with structural analyses might provide useful insight within this line of research.

In addition to HS, both HSPGs and enzymes participating in HS biosynthesis and editing appear to play a role in gastric carcinogenic events, as a few of these protein entities have been reported to be altered, and in some situations to correlate with patient's prognosis and survival.

In a study published by Wiksten *et al.*, SDC1 distinct expression patterns have been described in the epithelia and stroma of gastric adenocarcinoma samples. Loss of epithelial SDC1 was described in higher grade tumors and was associated with tumor location, larger tumor size, lymph node metastases, SDC1 detection in the stroma, tumor subtype, namely

the intestinal subtype, and patients' worse overall survival. Whereas stromal SDC1 high detection correlated with low epithelial SDC1, intestinal subtype tumors and patient's poor outcome (326). Overall, these results suggest that epithelial SDC1 might have a tumor suppressive role and /or that soluble SDC1 forms promote disease progression.

Also from the SDC family, SDC4 shows up as a unique molecule of interest in the gastric carcinogenesis pathway, determining gastric cancer aggressiveness and patients' clinical outcomes. Previously, Magalhães *et al.* reported a significant increase of SDC4 expression in the gastric mucosa of individuals infected by highly pathogenic strains of *H. pylori*, which is significantly associated with development of gastric pre-malignant lesions. In this work, it was further outlined a possible mechanism for this induction, englobing infection-mediated activation of the nuclear factor kappa B (NF-κB) transcriptional signaling and consequent increase of the production proinflammatory cytokine tumor necrosis factor alpha (TNF-α), which in its turn was shown to increase the expression of *SDC4* (327). More recently, our team has reported that SDC4 is indeed highly expressed on intestinal subtype gastric tumors, and it correlates with cancer aggressiveness features and patient worse prognosis. We also addressed the role of this glycoconjugate in cellular behavior and revealed that it promotes significantly both migration and invasion capabilities. Lastly, through *in vivo* assays we demonstrated that SDC4 is fundamental in defining EVs' tropism towards metastatic organs in gastric cancer, namely liver, lung and spleen (270).

Besides SDCs, other families of PGs have also been implicated in this malignancy, for instance the expression levels of GPC6 were found to be increased in gastric tumor tissues (328). Similarly, the extracellular CS/DS carrying SLRP biglycan was also found to be overexpressed, and it correlated with disease relapse and poor patient prognosis in advanced stages of disease (329).

Beyond these changes, several epigenetic and comparative genomic analyses reported deregulated expression of genes that code for different HS modifying enzymes in the context of gastric cancer. NDST1 was revealed to have a suppressive role in cell proliferation (330), upregulation of *HS6ST2* was associated with patients worse prognosis (331) and *HS3ST2* was found to be hypermethylated in gastric cancer (332). *HSULF1* promoter has also been reported to be hypermethylated in gastric cancer, resulting in *HSULF1* silencing (333). These results aligned with a posterior study revealing HSULF1 capability of suppressing cell proliferation, tumor growth and activity of the Wnt signaling pathway (334). Nonetheless, the contrary was also observed in a different study, reporting *HSULF1* upregulation and promoter hypomethylation. Regarding *HSULF2*, overexpression has been reported in this malignancy (335). Finally, *HPSE* high expression associated to gastric tumor growth and invasion and with patients poor prognosis (336).

Despite the current knowledge about the GAG and PG trends in the context of gastric cancer, discrepancies of the reported information and the lack of understanding about the molecular mechanisms underlying cell malignant behavior, make it harder to paint a clear picture of the GAGosylation landscape in gastric tumor microenvironment and how it contributes to tumor progression.

Overcoming the beforementioned general hurdles in this field, gaining further insight into the roles played by GAGs and GAGs enzymatic machinery in tumor cell transformation and proceeding with a reliable characterization of the GAG profile in disease, is fundamental to develop new diagnosis, management and treatment tools. The numerous studies describing the abnormal expression of HS and HSPGs in the context of cancer support the significant clinical potential these molecules hold as molecular biomarkers to improve cancer patients' stratification for efficient diagnosis and prognosis strategies. In addition, the development of pharmaceutical strategies that target HS and HS-protein interactions, including small molecule inhibitors of HS biosynthesis (337, 338), synthetic xylosides (339-341), HS mimetics (342) and anti-HS/HSPG antibodies (343, 344), has also been highlighted as a promising strategy to tackle cancer.

2. General Aim and Objectives

Gastric cancer, the fifth most commonly diagnosed cancer and fifth leading cause of cancer related death worldwide, persists as an aggressive and highly mortal pathology (2). Its asymptomatic nature throughout most of the disease progression phase and the current limited therapeutic strategies are main limitations that must be overcome, emphasizing the fundamental need to develop new sensitive diagnostic tools and efficient treatment options.

Aberrant glycosylation has been reported in the vast majority of cancer pathologies. Moreover, it has been implicated in multiple tumor cell deregulated mechanisms by influencing cell-cell and cell-ECM crosstalk as well as tumor microenvironment biophysical and biochemical properties. Particularly in the context of gastric cancer, a growing body of research indicates that changes in the expression of HSPGs and enzymes that modulate HS biosynthesis and post-synthesis editing, contribute to cell malignant transformation, tumor development and progression, and patient's unfavorable prognosis. These data support the clinical potential of HS and HS carriers to function as biomarkers for effective diagnosis and prognosis approaches, and as glyco-targets for new therapeutic solutions.

The general aim of this thesis is to understand the regulatory mechanisms underlying HS biosynthesis and editing in the context of gastric cancer, and to address the role of HS in the establishment of tumor cell aggressive phenotypes. To achieve this general aim, we focused on the following specific objectives:

Uncovering the role of HS glycosyltransferases EXTL2 and EXTL3 in gastric cancer, and the impact of aberrant HS biosynthesis in tumor cell behavior

Both in homeostasis and in cancer, the biological functions achieved by HSPGs in cell glycocalyx and ECM rely greatly on HS binding to a plethora of biologically active molecules. Although altered expression of HSPGs and HS biosynthesis enzymes has been described in several types of cancer, including gastric cancer, the specific implications of deregulated HS levels in tumor cell behavior remain unclear. In Chapter 2, we explored the roles of the HS glycosyltransferases EXTL2 and EXTL3 in fine-tuning HS and CS levels in gastric cancer, by establishing glycoengineered gastric cancer cell models and resorting to GAG disaccharide analyses. We assessed the impact of deregulated HS biosynthesis in cells' glycoproteome, and further addressed the biological function of HS in the activation of cell surface RTKs and in tumor cell motility and invasion. In addition, we assessed the contribution of the gastric cancer associated HSPG, SDC4, to gastric cancer cells' GAGosylation profile.

Addressing the role of HS 6-O-sulfation in the regulation of HS and HSPGs biosynthesis and cellular distribution in gastric cancer

Distinctive HS sulfation patterns are fundamental to determine the binding affinity of HSPGs towards specific ligand partners, allowing them to take over a range of tumor cell malignant events, by participating in several pro-oncogenic signaling cascades and influencing cell adhesion, migration and invasion. In Chapter 3, we focused on studying the modulation of HS 6-*O*-sulfation by HS-modifying and editing enzymes. We screened six gastric cancer cell lines, to evaluate the expression of these enzymes, and established glycoengineered gastric cancer cell lines with distinct HS sulfation profiles by targeting the 6-*O*-sulfotransferase HS6ST1 and 6-*O*-endosulfatase HSULF2. We studied the regulatory roles of these enzymes in cellular GAGosylation using GAG disaccharide structural analysis, transcriptomics and a 6-*O*-endosulfatase enzymatic assay. We further addressed the clinical relevance of HS 6-*O*-sulfation epitopes in gastric cancer by analyzing the transcriptomic profile of *HS6ST1* and *HSULF2* in samples from a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD).

Identification of a gastric cancer-associated GAGosylation signature

Various studies have tried to characterize the GAGs expression profile on stomach tumors and have reported changes in the expression of several HSPGs and enzymes that participate in HS and CS biosynthesis. Nonetheless, it remains to be identified a clear distinctive GAGosylation feature that could be applied in cancer therapy. In Chapter 4, we aimed to uncover distinctive gastric cancer-associated GAGosylation profiles, by screening a set of gastric cancer cell lines, with different histopathological origins, and evaluating the expression of HS- and CS-related enzymes and HSPGs. Moreover, TCGA-STAD data mining and bioinformatic analysis were also performed to identify specific tumor-associated HS and CS/DS transcriptomic profiles, which were also correlated with patients' prognosis.

Chapter 2

Uncovering the role of the glycosyltransferases Exostosin-Like 2 (EXTL2) and EXTL3 in gastric cancer, and the impact of aberrant Heparan Sulfate levels in tumor cell motility and invasion

Glycosyltransferases EXTL2 and EXTL3 cellular balance dictates heparan sulfate biosynthesis and shapes gastric cancer cell motility and invasion

Catarina Marques[1,2,3], Juliana Poças[1,2,4], Catarina Gomes[1,2], Isabel Faria-Ramos[1,2], Celso A. Reis[1,2,4,5], Romain R. Vivès[6], Ana Magalhães[1,2,4*]

[1] i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal

[2] IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto, 4200-465 Porto, Portugal

[3] Programa Doutoral em Biologia Molecular e Celular (MCbiology), Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal.

[4] ICBAS - Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, 4050-313 Porto, Portugal

[5] FMUP - Faculdade de Medicina da Universidade do Porto, 4200-319 Porto, Portugal

[6] Univ. Grenoble Alpes, CNRS, CEA, IBS, Grenoble, France

[*] Corresponding author: amagalhaes@ipatimup.pt

*Journal of Biological Chemistry (*2022) doi: 10.1016/j.jbc.2022.102546.

ABSTRACT

Heparan Sulfate (HS) Proteoglycans (HSPGs) are abundant glycoconjugates in cells' glycocalyx and Extracellular Matrix. By acting as scaffolds for protein-protein interactions, HSPGs modulate extracellular ligand gradients, cell signaling networks, and cell-extracellular matrix crosstalk. Aberrant expression of HSPGs and enzymes involved in HSPG biosynthesis and processing has been reported in tumors, with impact in cancer cell behavior and tumor microenvironment properties. However, the roles of specific glycosyltransferases in the deregulated biosynthesis of HSPGs are not fully understood. In this study, we established glycoengineered gastric cancer cell models lacking either Exostosin-Like glycosyltransferase 2 (EXTL2) or EXTL3 and revealed their regulatory roles in both HS and Chondroitin Sulfate (CS) biosynthesis and structural features. We showed that EXTL3 is key for initiating the synthesis of HS chains in detriment of CS biosynthesis, intervening in the fine-tuned balance of the HS/CS ratio in cells, while EXTL2 functions as a negative regulator of HS biosynthesis, with impact over the glycoproteome of gastric cancer cells. We demonstrated that KO of *EXTL2* enhanced HS levels along with concomitant upregulation of Syndecan-4, which is a major cell surface carrier of HS. This aberrant HS expression profile promoted a more aggressive phenotype, characterized by higher cellular motility and invasion, and impaired activation of Ephrin type-A 4 cell surface receptor tyrosine kinase. Our findings uncover the biosynthetic roles of EXTL2 and EXTL3 in the regulation of cancer cell GAGosylation and proteoglycans expression and unravel the functional consequences of aberrant HS/CS balance in cellular malignant features.

INTRODUCTION

Heparan Sulfate (HS) Proteoglycans (HSPGs) comprise an abundant class of glycoconjugates, composed by a core protein with covalently attached HS glycosaminoglycan (GAG) chains. These molecules are ubiquitously expressed, both at the cell surface and in the Extracellular Matrix (ECM), and play key roles in cellular physiology, impacting cellular proliferation, adhesion and motility, membrane trafficking, formation of extracellular gradients, morphogenesis, and angiogenesis (86). HSPGs are also key players in pathological scenarios, being described as important maestros of cancer cell interaction with the ECM, regulating cancer cell communication, and modulating the tumor microenvironment with impact in disease progression (5). HS chains present high binding affinity to multiple biologically active partners, including transmembrane receptors, ECM structural proteins, and soluble molecules. Therefore, HS chain glycan composition dictates the biological activities of HSPGs (345-347).

HS chains are linear unbranched polysaccharides composed by repeating disaccharide units of glucosamine and uronic acid residues (4). HS biosynthesis occurs in the Golgi

apparatus or at the endoplasmic reticulum-Golgi interface and is organized in three major events: (i) GAG-protein linker assembly, (ii) HS chain polymerization, and (iii) structural modifications of the elongated chain (90). The linker assembly is initiated by the addition of a xylose (Xyl) residue to specific serine (Ser) residues on the protein core (348). This is followed by the sequential addition of two galactose (Gal) and one glucuronic acid (GlcA) residues that form the universal tetrasaccharide GAG linker GlcAβ1-3Galβ1-3Galβ1-4Xylβ1-*O*-Ser, through which HS chains are covalently attached to the protein moiety (4, 349). The transfer of each monosaccharide residue is catalyzed by a particular glycosyltransferase. The assembly of this GAG-linker also involves additional modification steps, including the transient phosphorylation of the Xyl residue by the kinase FAM20B and phosphatase XYLP, after the addition of the first Gal residue. This step enhances the activity of subsequent acting glycosyltransferases and promotes the maturation of the linker (140, 141, 144). The addition of a single *N*-acetylglucosamine (GlcNAc) residue to the tetrasaccharide linker then initiates the polymerization of the HS chains, whereas addition of a *N*-acetylgalctosamine (GalNAc) residue would orientate the biosynthesis pathway towards the assembly of Chondroitin Sulfate (CS)/Dermatan Sulfate (DS) chains. Addition of this first GlcNAc is catalyzed by two members of the Exostosin (EXT) family, Exostosin-Like 2 (EXTL2) and EXTL3, and is followed by further elongation promoted by a hetero-oligomeric complex formed by EXT1 and EXT2, which will catalyze the alternated transfer of GlcNAc and GlcA residues (149-151, 350). Once polymerized, the pro-heparan chains undergo extensive processing by HS modifying enzymes, including GlcNAc deacetylation and sulfation by *N*-Deacetylase/*N*-Sulfotransferases (NDST1-4) (351), epimerization of GlcA into iduronic acid (IdoA) by Glucuronyl C5-epimerase (157), and sulfation at multiple positions by different *O*-Sulfotransferases (HS2ST1, HS6ST1-3 and HS3ST1-6), leading to the synthesis of mature HS chains displaying highly variable structures and functions (135, 163, 228). These HS chains can then be further modified postsynthetically by heparanase cleavage and 6-*O*-desulfation catalyzed by extracellular 6-*O*-endosulfatases (Sulfs), thus generating additional structural diversity with biological relevance (90).

The members of the Exostosin family that dictate the initiation of HS chain biosynthesis, EXTL2 and EXTL3, have been studied regarding their impact on GAG content using both *in vitro* and *in vivo* models. However, the molecular mechanisms underlying the HS elongation steps are not fully understood yet (352). It has been demonstrated that EXTL3 catalyzes with high efficiency the transfer of the first GlcNAc residue to the mature tetrasaccharide linker and that it can also participate in HS elongation by adding GlcNAc to the growing chain, while being inefficient in the transfer of GlcA (149, 353). Gene KO experiments performed in mouse models revealed that systemic loss of EXTL3 expression led to early embryonic lethality, furthermore the inactivation of *EXTL3* specifically in mice

pancreatic islet beta cells caused impaired HS biosynthesis (354), which was also observed in *EXTL3* zebrafish mutants (355). In the same line, more recently, it was shown that KO of *EXTL3* in Chinese hamster ovary (CHO) cells resulted in the abolition of HS expression, which further indicates the crucial role of this enzyme in initiating the biosynthesis of HS chains (356). These results are in agreement with earlier observations that *EXTL3* silencing led to longer HS chains, most probably because of the reduced number of chains being synthesized (357).

There is still reasonable doubt concerning the role of EXTL2 in HS biosynthesis regulation. EXTL2 is an α1,4-N-acetylhexosaminyltransferase, exhibiting dual *in vitro* catalytic activity. Enzymatic assays have shown that this enzyme can act both as a α-GlcNAc and as a α-GalNAc glycosyltransferase towards synthetic linker mimetics. However, it was demonstrated that EXTL2 could not add GlcNAc residues to a mature tetrasaccharide linker substrate (GlcAβ1–3Galβ1–3Galβ1-4Xylβ1-*O*-Ser) (151). More recently, it has been proposed that EXTL2 could mediate an alternative FAM20B-dependent pathway that suppresses HS biosynthesis. According to this model, in the initial steps of HS/CS formation, which entail linker assembly, increased FAM20B kinase activity and/or reduced dephosphorylation by XYLP promote formation and accumulation of the phosphorylated tetrasaccharide linker (GlcAβ1-3Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)). This phosphorylated sequence can be used as substrate by EXTL2, that catalyzes the transfer of a GlcNAc residue and generates an immature phosphorylated pentasaccharide (GlcNAcβ1–4GlcAβ1-3Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)) that cannot be further polymerized by EXT1 and EXT2, thus resulting in the premature termination of HS elongation (144, 148). This model is in agreement with the increment of HS disaccharides observed on CHO *EXTL2* KO cells (356), as well as in *EXTL2* deficient mice (148, 358). However, the regulatory role played by this enzyme in HS biosynthesis still remains controversial. Whereas the downregulation of EXTL2 resulted in increased HS chains length in human embryonic kidney 293 (HEK293) cells, no significant changes were detected when these cells overexpress EXTL2 (359).

The deregulation of HS biosynthetic machinery has been described as an important event underlying HS abnormal expression in cancer (5). Colorectal (249, 286, 287), breast (240, 285), lung (304), hepatocellular (247), and gastric (327, 330-332) carcinoma are some of the malignant conditions where altered expression of HSPGs and HS-related genes has been reported. Particularly, gastric tumors have been shown to display aberrant GAGs overall content and altered sulfation patterns (208, 325). However, the mechanisms underlying the alteration of GAG biosynthesis and HS structure in cancer still need to be further clarified. Additionally, the cell- and tissue-specific GAG biosynthesis regulation by GAG-related enzymes, whose expression is variable at the different stages of pathologies,

underlines the importance of addressing the impact of these glycosyltransferases' enzymatic activities on cancer cell behavior and tumor microenvironment remodeling.

In this work, we have evaluated the roles of EXTL2 and EXTL3 in HS biosynthesis in gastric cancer. In addition, we determined the consequences of changes in their expression on the structural features of GAGs in tumor cells' glycocalyx and their functional impact in cancer cell aggressiveness.

RESULTS

EXTL2 and EXTL3 expression regulates the cellular HS content

To investigate the functional roles of EXTL2 and EXTL3 in HS biosynthesis, KO cell models were generated from the gastric cancer cell line MKN74 *via* CRISPR-Cas 9 genome editing (Fig. S1A and B). We first evaluated the effects of *EXTL2* or *EXTL3* gene KO on HS cellular content. Flow cytometry analysis was performed to measure cell surface HS and revealed that in the absence of EXTL2, cells show increased levels of HS (Fig. 1A). Immunofluorescence data showed HS staining in cells' plasma membrane and cytoplasm, for both WT and *EXTL2* KO cells, though a higher number of *EXTL2* KO cells were positive for 10E4 (Fig. 1B). In line with these observations, Western blot (WB) assays revealed increased total HS content in *EXTL2* KO cells, as shown by the stronger HS signal covering a wide range of molecular weight (MW), corresponding to the highly variable HS structures decorating HSPGs (Fig. 1C). In contrast, KO of *EXTL3* led to the complete depletion of HS, both at the cell surface (Fig. 1A) and total cell content (Fig. 1B and C). To further investigate the role of EXTL2 and EXTL3 in the regulation of HS biosynthesis, we examined the number of HS chains in our glycoengineered cell models. Cells were treated with Heparinase III (Hep. III) that specifically digests HS chains and generates smaller HS stubs capped with unsaturated GlcA residues, which are recognized by the 3G10 antibody. Protein extracts from *EXTL2* KO cells showed stronger 3G10 labelling, which was particularly evident around 37 kDa (Fig. 1D), suggesting an increase in the number of HS chains upon abrogation of EXTL2. *EXTL3* KO cells lacked these HS stub structures (Fig. 1D), in accordance with the data shown in Fig. 1A- C.

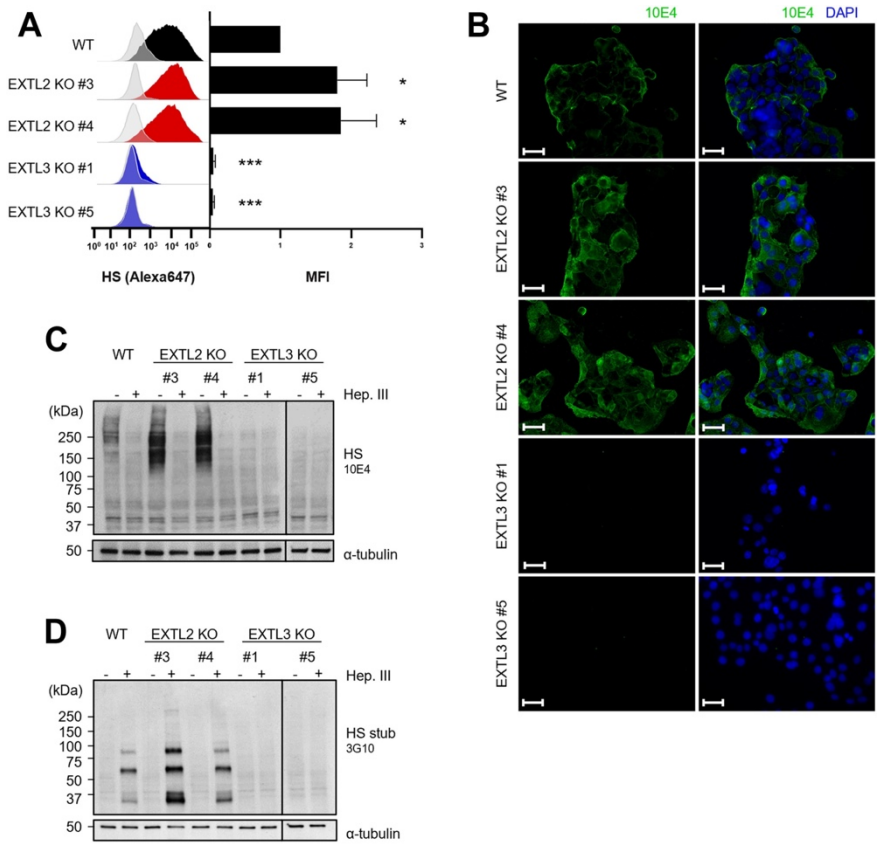


Figure 1. *EXTL2* KO and *EXTL3* KO gastric cancer cells display altered cell surface and overall expression of HS. (A) The levels of cell surface HS in MKN74 WT, *EXTL2* KO and *EXTL3* KO cells were quantified by flow cytometry analysis. *Gray* shaded peaks indicate the negative controls. The peaks colored in *solid black*, *red* or *blue* indicate the labelling with 10E4 mAb of WT, *EXTL2* KO or *EXTL3* KO cells, respectively. Bar graphs show normalized mean intensity of fluorescence (MIF) + SD. n=3 independent biological experiments **(B)** Immunofluorescence labelling of HS 10E4 epitopes (*green*) was performed in MeOH fixed cells. Nuclei (*blue*) were stained with DAPI. Scale bar 50 µm. **(C and D)** MKN74 WT, *EXTL2* KO and *EXTL3* KO cells were evaluated for HS total content and number of HS chains by WB analyses. Protein lysates were treated with or without Heparinase III (Hep. III) and immunoblotted for undigested HS using the 10E4 mAb **(C)** or HS stubs using the 3G10 mAb **(D).** Images are representative of 2 independent experiments.

EXTL2 and EXTL3 impact Syndecan-4 expression and its glycosylation profiles

Taking in consideration the impact on HS biosynthesis observed in the glycoengineered cell models and knowing that Syndecans (SDCs) and glycosylphosphatidylinositol-anchored (GPI)-anchored Glypicans (GPCs) are major carriers of HS on epithelial cells, we then evaluated the impact of *EXTL2* and *EXTL3* KO over SDCs and GPCs expression in the gastric cancer cells.

Transcriptomic analyses revealed that the expression of *GPC1* and *GPC5* was not altered in glycoengineered models, *GPC4* showed lower expression in *EXTL2* KO and *EXTL3* KO, though the expression level in the WT was already very low, and *GPC2*, *GPC3* and *GPC6* were not detected (Fig. 2A). Regarding SDC family members, *SDC1* expression remained unaltered upon KO of *EXTL2* and *EXTL3*, *SDC2* was not detected, and we observed a trend towards *SDC3* overexpression in the glycoengineered models (Fig. 2A). Noteworthy, the expression of *SDC4* was significantly increased in both *EXTL2* KO clones,

while for *EXTL3* KO, the expression was variable as the increase was only significant for one of the clones (Fig. 2A).

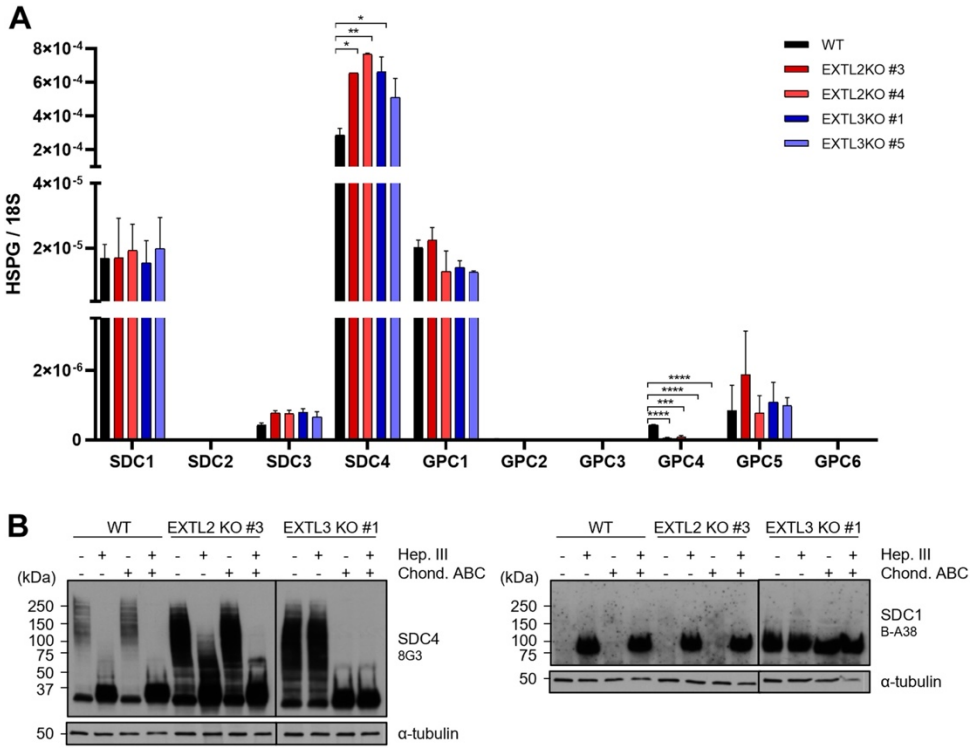


Figure 2. Modulation of HSPGs' expression and GAGosylation profiles by *EXTL2* and *EXTL3* KO cells. (**A**) The expression of cell surface HSPGs, SDCs 1-4 and GPCs 1-6, was assessed in MKN74 KO cell models by qRT-PCR. Bar graph shows the mean RQ values for each gene, normalized to the internal control gene 18S + SD. Two independent biological assays with technical triplicates were analyzed. (**B**) The protein levels and GAGosylation profiles of SDC4 and SDC1 in *EXTL2* KO and *EXTL3* KO cells were evaluated by WB analysis. Protein lysates were treated with or without Heparinase III (Hep. III) and/or Chondroitinase ABC (Chond. ABC) and immunoblotted for SDC4 and SDC1 using 8G3 and B-A38 mAb, respectively. Images are representative of 2 independent experiments.

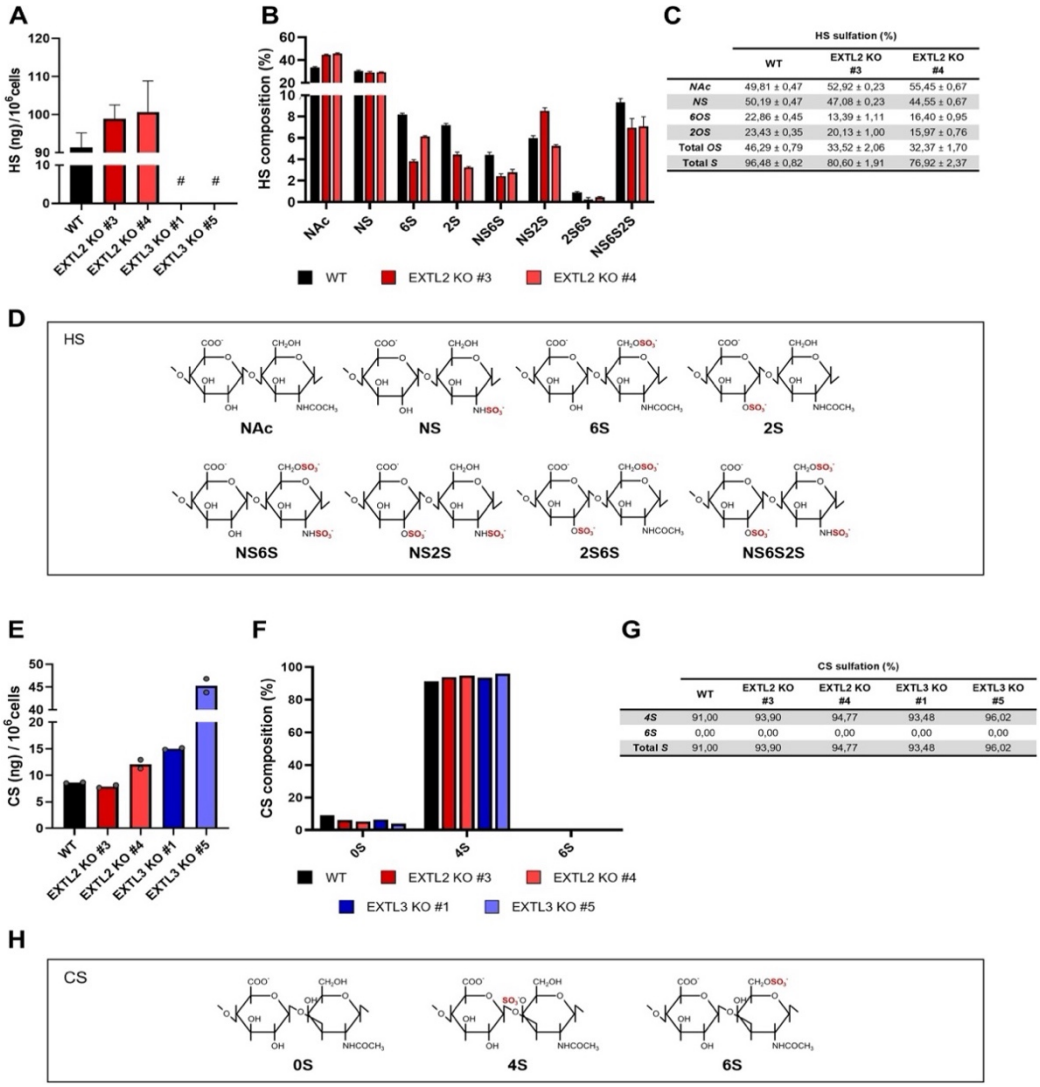
Since we observed that MKN74 cells express mainly the SDC1 and SDC4, we further evaluated their protein levels and GAGosylation profiles by WB. Higher SDC4 protein levels were found in the *EXTL2* KO cell model, as evidenced by the increased staining of the fully glycosylated (smear staining) and unglycosylated (37 kDa band staining) SDC4 molecules observed in the non-digested cell lysates (Fig. 2B). Cleavage of HS chains by Hep. III on WT cell lysates led to a marked shift in the staining profile of SDC4, from a smear that covered high MW to a sharp staining at 37 kDa (deglycosylated SDC4). In *EXTL2* KO cell lysates upon Hep. III digestion, an intense smear could still be detected in the lower MW range, which was considerably shortened upon combined digestion with Chondroitinase ABC (Chond. ABC) (Fig. 2B). These results suggest that in the absence of *EXTL2* gene expression, SDC4 might also be modified with CS/DS chains, and that Chond. ABC enzymatic activity is more efficient upon HS removal. This could explain the differences in SDC4 labelling when comparing Hep. III single digestion and Hep. III + Chond. ABC double

digestion, as well as the similar staining between nondigestion and Chond. ABC single digestion (Fig. 2B). WB analyses of *EXTL3* KO nondigested cell lysates revealed increased signal for glycosylated SDC4, with a large smear between 50 kDa and 250 kDa, while no variations were observed for the labelling of the unglycosylated form at 37 kDa (Fig. 2B). This suggests that HS chains might constitute hindrance to the binding of the anti-SDC4 8G3 antibody to the core protein extracellular domain, hence the increased labelling of glycosylated SDC4 in the cell model that lacks these chains. No significant variations were detected in the labelling of SDC4 in *EXTL3* KO cells upon single digestion with Hep. III, in agreement with the previous results showing lack of HS in this cell model (Fig. 1). Interestingly, CS/DS cleavage by Chond. ABC eliminated the SDC4 labelling smear (Fig. 2B), which further supports that alternative GAGosylation may occur and that the SDC4 expressed in *EXTL3* KO cells is modified with CS/DS chains, instead of HS. Regarding SDC1, HS digestion revealed similar labelling in the WT and *EXTL2* KO cells, which was only detected in Hep. III digested samples, while in *EXTL3* KO cells that lack HS, GAG digestion had no impact over SDC1 labelling profile (Fig. 2B).

EXTL2 and EXTL3 abrogation changes HS and CS GAG structural motifs in gastric cancer cells

To characterize HS and CS glycan structures from *EXTL2* KO and *EXTL3* KO glycoengineered cancer cell lines, we isolated and purified cellular GAGs and performed HS and CS disaccharide analyses using reversed-phase ion pair HPLC (RPIP-HPLC).

Structural analyses showed a trend for increased HS disaccharide content in *EXTL2* KO cells, in comparison with WT, while in *EXTL3* KO cells HS disaccharides were not detected (Fig. 3A). Also, HS composition appeared to vary between WT and *EXTL2* KO cells (Fig. 3B and C). Results suggest that in the absence of EXTL2 expression, HS exhibits increased levels of nonmodified *N*-acetylated disaccharides, along with a decrease in overall disaccharide *O*-sulfation (Fig. 3B-D). Regarding CS composition, *EXTL3* KO resulted in increased CS disaccharide content (Fig. 3E). This observation is in agreement with the high MW smear observed for SDC4 in the *EXTL3* KO clones, which was abolished after Chond. ABC digestion (Fig. 2B). The effect of *EXTL2* KO on CS expression was more heterogeneous among clones, with only one presenting increased CS disaccharide units (Fig. 3E). Furthermore, no major variations in CS composition were detected for MKN74 glycoengineered cells compared to the WT (Fig. 3F-H).



HS sulfation (%)	WT	EXTL2 KO #3	EXTL2 KO #4
NAc	49,81 ± 0,47	52,92 ± 0,23	55,45 ± 0,67
NS	50,19 ± 0,47	47,08 ± 0,23	44,55 ± 0,67
6OS	22,86 ± 0,45	13,39 ± 1,11	16,40 ± 0,95
2OS	23,43 ± 0,35	20,13 ± 1,00	15,97 ± 0,76
Total *OS*	46,29 ± 0,79	33,52 ± 2,06	32,37 ± 1,70
Total *S*	96,48 ± 0,82	80,60 ± 1,91	76,92 ± 2,37

CS sulfation (%)	WT	EXTL2 KO #3	EXTL2 KO #4	EXTL3 KO #1	EXTL3 KO #5
4S	91,00	93,90	94,77	93,48	96,02
6S	0,00	0,00	0,00	0,00	0,00
Total *S*	91,00	93,90	94,77	93,48	96,02

Figure 3. HS and CS disaccharide composition in *EXTL2* KO and *EXTL3* KO gastric cancer cells. (A) HS chains isolated and purified from MKN74 cells were digested with a mixture of Heparinases I, II and III and analyzed by RPIP-HPLC to determine the total amounts and structural composition of HS disaccharides in each cell model. Bar graphs show the mean levels of HS disaccharides, expressed in ng. # represents the conditions where HS disaccharides were non-detected. Technical triplicates were analyzed for the WT and *EXTL2* KO models, and technical duplicates were analyzed for *EXTL3* KO models. The analyzed samples reflect pooling from 2 independent biological replicates. **(B)** Relative quantities of distinct HS disaccharide units were measured. Bar graph shows the mean levels of each type of HS disaccharide: NAc = ΔHexUA-GlcNAc; NS = ΔHexUA-GlcNS; 6S = ΔHexUA-GlcNAc,6S; 2S = ΔHexUA,2S-GlcNAc; NS6S = ΔHexUA-GlcNS6S; NS2S = ΔHexUA,2S-GlcNS; 2S6S = ΔHexUA,2S-GlcNAc,6S; NS2S6S = ΔHexUA,2S-GlcNS6S. (**C**) Table represents the mean values of sulfation content. *NAc* = N-Acetylation; *NS* = *N*-Sulfation; *6S* = 6-*O*-Sulfation; *2S* = 2-*O*-Sulfation; total *OS* = total *O*-Sulfation, total *S* = total Sulfation (*N*- and *O*-sulfation). (**D**) Illustrative representation of HS disaccharide structures is included. (**E**) CS chains isolated and purified from MKN74 cells were digested with Chondroitinase ABC and were also analyzed by RPIP-HPLC. Bar graph shows relative mean quantities of distinct CS disaccharide units. Technical duplicates were analyzed from the pool of 2 independent biological samples. (**F**) Relative quantities of distinct CS disaccharide units were measured. Bar graph shows the levels of each type of CS disaccharide: 0S = ΔHexUA-GalNAc 4S = ΔhexUA-GalNAc,4S; 6S = ΔhexUA-GalNAc,6S. One dataset is represented. (**G**) Table represents the CS sulfation content in each cell model. *4S* = 4-*O*-Sulfation; *6S* = 6-*O*-Sulfation; total *S* = total Sulfation (4-*O*- and 6-*O*-Sulfation). (**H**) Illustrative representation of CS disaccharide structures is included.

Aberrant GAGosylation profiles displayed by *EXTL2* KO cancer cells promote motile and invasive phenotypes

GAGs are key modulators of ECM physical and biochemical properties, which are capable of binding to different proteins and allow PGs to act as cell surface co-receptors and mechanosignaling transducers, ultimately impacting cell behavior (94). Taking these biological features in consideration, we resorted to our glycoengineered cell models to disclose the impact of the distinct GAGosylation profiles on the migration and invasion capabilities of gastric cancer cells.

We first assessed whether *EXTL2* or *EXTL3* KO would impact cells viability and proliferation and found that abrogation of EXTL2 or EXTL3 had no impact, neither on the percentage of viable cells (Q4, Fig. 4A) nor on the cells' proliferating percentage (Fig. 4B). To determine the functional impact of altered GAGosylation in cancer cell motility, we performed wound healing assays and evaluated the migration rates of *EXTL2* KO and *EXTL3* KO cells on fibronectin (Fig. 4C) and collagen IV (Fig. 4D) coated surfaces. Interestingly, the *EXTL2* KO cells presented higher migration rates in both coatings, while no significant differences were observed for *EXTL3* KO cells when compared to the WT (Fig. 4C and D). All cells migrated faster in the presence of collagen IV than in fibronectin.

In order to assess if *EXTL2* KO would also impact cancer cell invasion, we evaluated cells' ability to break down and penetrate matrigel-coated membranes, which is a basement membrane-like matrix, enriched in ECM proteins. Remarkably, *EXTL2* KO cells displayed significantly higher invasion capabilities, while *EXTL3* KO cells showed no differences compared to the WT (Fig. 4E). These results indicate that *EXTL2* KO cells' aberrant display of GAGs promotes higher aggressiveness features such as more motile and invasive phenotypes, which are not dependent on altered viability or proliferation.

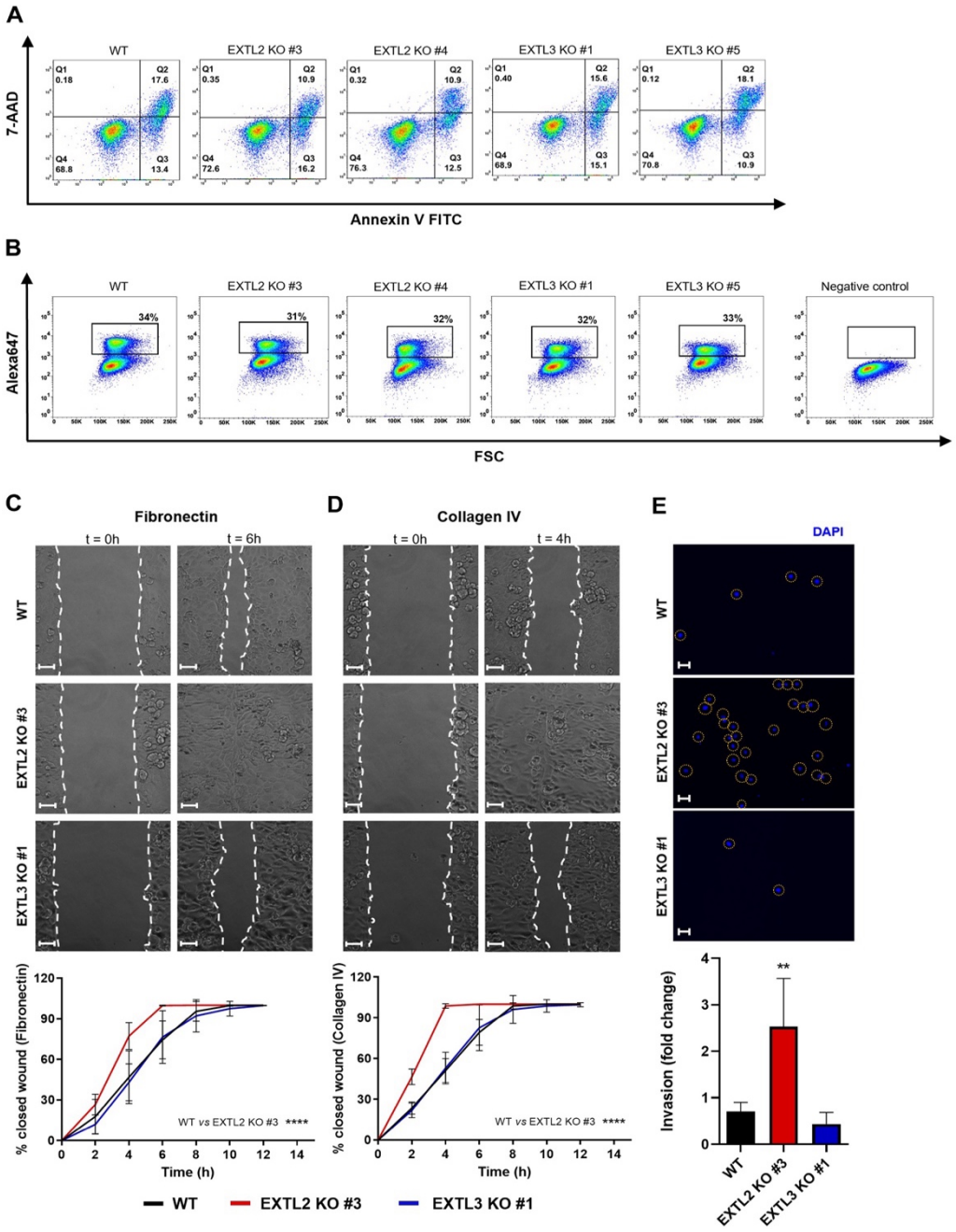


Figure 4. Impact of EXTL2 and EXTL3 expression in MKN74 gastric cancer cell biological features (A) Cell viability was evaluated by performing 7-AAD/ Annexin V-FITC cell labelling. Dot plots show percentages of viable cells (Q4), cells in early apoptosis (Q3) cells in late apoptosis (Q2) and cells in necrosis (Q1). Data from one representative assay from two independent biological replicates. **(B)** Cell proliferation was assessed by measuring EdU incorporation in cell's DNA by flow cytometry. Dot plots show the percentage of proliferating cells as EdU-Alexa647 positive cells, highlighted in the *black* box. A representative dot plot of a negative control (without EdU) is shown. Data from one representative assay from two independent biological replicates. **(C and D)** Migration capabilities of MKN74 WT and one representative clone of the *EXTL2* KO and *EXTL3* KO cell models were addressed by performing a wound healing migration assay. Representative pictures of the wounded cell monolayer in fibronectin **(C)** and collagen IV **(D)** coated slides are depicted together with the graphs showing the mean percentages of closed wounds + SD over time. Scale bar 50 µm. Two independent biological experiments with technical triplicates were analyzed. **(E)** Invasion was evaluated by performing a Matrigel invasion assay. Microscope images show highlighted DAPI stained nuclei of the invasive cells counted for each condition. Scale bar 50 µm. Bar graph shows the mean percentage of invading cells + SD. Two independent biological experiments with technical triplicates were analyzed.

GAGs remodeling by *EXTL2* KO impacts Ephrin type-A receptor 4 activation

Receptors tyrosine kinase (RTKs) activation and downstream signaling pathways are known to promote aberrant cellular events related to the acquisition of cancer hallmark capabilities, which contribute to tumor progression. GAG chains, in particular HS, are able to modulate cancer cell signaling by coupling multiple biologically active ligands, like growth factors, to their targeted RTKs. This prompts HSPGs to act as important scaffolds for protein–protein interactions and to trigger receptors activation and subsequent signaling transduction (38).

To uncover the impact of specific GAGosylation profiles from *EXTL2* KO and *EXTL3* KO gastric cancer cells in the activation of RTKs, we screened the phosphorylation state of multiple RTKs using a phospho array (Fig. S2A and B). Results from the phospho-RTK array indicated that *EXTL2* KO cells exhibited higher insulin-like growth factor 1 receptor (IGF-I R) and epidermal growth factor receptor (EGFR) activation, and lower Ephrin type-A receptor 4 (EphA 4) phosphorylation when compared to the WT. Interestingly, *EXTL3* KO cells showed increased EGFR and IGF-IR phosphorylation, similar to *EXTL2* KO, but higher EphA 4 activation (Fig. 5A). Overall, the RTK-array revealed EphA 4 as an interesting candidate to be further studied since *EXTL2* KO and *EXTL3* KO cells presented an opposite trend in terms of its activation.

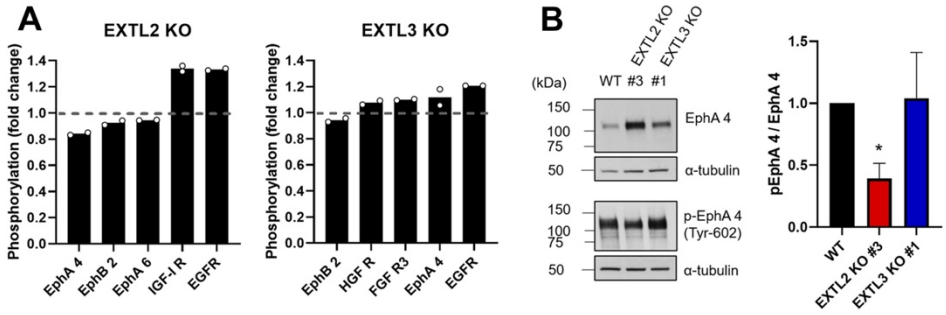


Figure 5. Abrogation of EXTL2 expression decreases EphA 4 phosphorylation in MKN74 gastric cancer cells. (A) Phospho-RTKs in MKN74 glycoengineered cell models were analyzed. Bar graphs show the mean relative phosphorylation of different RTKs expressed in *EXTL2* KO and *EXTL3* KO cells normalized to WT values (WT=1). For each representative KO clone (*EXTL2* KO #3 and *EXTL3* KO #1), only RTKs whose phosphorylation fold change varied over 0,05 compared to the WT are shown. **(B)** The activation of EphA 4 was measured by WB analysis. α-Tubulin was used as loading control. Bar graph shows the mean values of the normalized phosphorylated receptor + SD. n=3 independent biological experiments.

Validation analysis by WB for total EphA 4 receptor amounts and phosphorylation at Tyr-602 revealed that abolishing the expression of EXTL2 in gastric cancer cells, which is concomitant with higher HS levels, resulted in increased levels of the total receptor with a statistically significant reduction on its activation (Fig. 5B). This result supports a possible role of HS remodeling by *EXTL2* KO in cellular signaling networks mediated by EphA 4. The validation analysis of EphA 4 phosphorylation at Tyr-602 in *EXTL3* KO cells, did not show

statistically significant differences in comparison with WT cells (Fig. 5B). Regarding EGFR activation, we analyzed by WB the specific phosphorylation of Tyr-1068 and Tyr-1086 in the different clones and could not confirm increased phosphorylation at these particular Tyr residues (Fig. S2C).

DISCUSSION

HSPGs are important components of cells' glycocalyx and ECM. By binding to multiple biological ligands *via* their sulfated HS chains, HSPGs modulate cancer cells' interactions with ECM and signaling networks, ultimately controlling tumor microenvironment and disease progression (5). In the present study, we disclosed the regulatory roles of EXTL2 and EXTL3 in GAG and HSPG biosynthetic pathways, in the context of gastric cancer, and evaluated the impact of the resulting altered GAGosylation in cancer cells' motility and signaling features.

For this purpose, we have established glycoengineered gastric cancer cell models by knocking-out either *EXTL2* or *EXTL3* from MKN74 cells and evaluated the impact of both glycosyltransferase KOs over GAG biosynthesis. *EXTL2* KO led to increased cell levels of HS, while *EXTL3* KO fully abolished HS expression (Fig. 1 and Fig. 3A). These data demonstrate the role of EXTL2 as a negative regulator of HS biosynthesis and further corroborate EXTL3-mediated promotion of HS polymerization. These results are in agreement with previous observations in different cellular (356) and animal (148, 354, 355, 358) models. Interestingly, a previous report has revealed that HS biosynthesis could be driven by EXTL2 and EXT2 in the context of EXT1-deficient mouse L fibroblasts, whereas in the same model EXTL3 silencing had little effect on cellular HS levels (360). Together with our data, these observations suggest that the phosphorylated pentasaccharide linker, formed after the transfer of a GlcNAc residue by EXTL2 to the phosphorylated tetrasaccharide, could serve as substrate for EXT2, whereas usually it is not further polymerized by the heterodimeric complex EXT1-EXT2 (148). In contrast, complete abrogation of HS biosynthesis upon KO of *Ext1* and/or *Ext2* was reported in CHO cells (356). Overall, these results support that these glycosyltransferases act in a tissue- and cell-specific manner. Therefore, it remains critical to further study the complex interactions of GAG biosynthetic machinery elements and its impact on the regulatory mechanisms underlying GAG biosynthesis in different cellular contexts and disease.

We investigated the impact of deregulated HS biosynthesis in the expression and GAGosylation features of two major HS carriers, SDC1 and SDC4. Our data showed that SDC4 was significantly overexpressed in *EXTL2* KO cells (Fig. 2A and B), which indicates that HS chains, beside impacting HSPG cellular functions, might also intervene in the regulation of cell membrane HSPG expression and therefore in the remodeling of cell

glycoproteome. The enhanced SDC4 expression could also explain the increased number of HS chains detected for this particular cell model (Fig. 1D). Analysis of protein levels also showed increased SDC4 detection on *EXTL3* KO cells (Fig. 2B), which was not reflected at the mRNA levels for both clones (Fig. 2A). Therefore, we hypothesize that 8G3 antibody has higher binding affinity to its target epitopes in HS-free SDC4 expressed in *EXTL3* KO cells. This may also explain the higher signal intensity observed for the 8G3-labelled de-GAGosylated SDC4 in the WT. Regarding *EXTL2* KO cells, our data suggest that SDC4 overexpression, and consequent higher levels of available 8G3 binding epitopes, overcomes the effects of HS-hindrance (Fig. 2B). It has been previously reported that the endoglycosidase heparanase influences the cellular distribution of SDC1 and SDC4 and plays a role in HSPG turnover (361). In the same study, not only HSPG core protein, but also HS chains were found to be critical for the efficient HSPG-heparanase binding and its cellular uptake, whereas other GAGs like CS could not perform this task (361). Therefore, we can also hypothesize that the aberrant GAGosylation profile of SDC4, resulting from *EXTL3* KO, might impact internalization, degradation and/or recycling of this HSPG, which could ultimately lead to its abnormal cellular accumulation.

Regarding HSPG GAGosylation profiles, GAG enzymatic digestion assays revealed that SDC4, that normally carries only HS, was abnormally modified with CS chains in the absence of *EXTL3* expression (Fig. 2B), which was further corroborated by CS structural analysis that showed increased CS disaccharide levels in *EXTL3* KO cell models (Fig. 3E). Curiously, Gopal *et al.* have previously shown that deregulation of HS biosynthesis by mutagenesis of SDC4 GAG attachment sites affected HS/CS ratio with increased CS levels (362). Furthermore, accumulation of CS was also reported in CHO cell mutants exhibiting impaired biosynthesis of HS chains (356, 363). In line with these observations, significant higher levels of CS were found on *extl3* mutant zebrafish larvae, and a larger proportion of CS chains were specifically attached to SDC4 upon silencing of *EXTL3* in HEK293 (364). Our results support that abrogation of EXTL3 favors the accumulation of available tetrasaccharide linker substrates, attached to specific HSPGs, which cannot be polymerized by EXT1 and EXT2 and are therefore available for priming CS assembly by Chondroitin *N*-acetylgalactosaminyltransferases. However, the rationale for HSPG glycosylation profile changes seen in *EXTL2* KO cell models (Fig. 2B) might be different. Izumikawa T. *et al.* have shown that after linker assembly, the CS polymerase Chondroitin *N*-acetylgalactosaminyltransferase 1 adds preferably a GalNAc residue to the extremity of the phosphorylated tetrasaccharide linker, generating the pentasaccharide GalNAcβ1-4GlcAβ1-3Galβ1-3Galβ1-4Xyl(2-*O*-phosphate). This reaction is accompanied by the dephosphorylation of the linker by XYLP, and followed by CS elongation catalyzed by Chondroitin (Chn) polymerases (365). Similarly, we hypothesize that in our *EXTL2* KO cell

models, lack of EXTL2 expression in MKN74 gastric cancer cells could lead to the accumulation of phosphorylated tetrasaccharides that are not capped with GlcNAc residues and are thus available to undergo this particular CS biosynthesis pathway. At the same time, EXTL3-driven HS biosynthesis could be conserved (Fig. 2B). Moreover, the changes observed in the labelling profile of SDC4 upon double digestion with Hep. III + Chond. ABC, when compared with single digestion with Hep. III, further support modification of SDC4 with CS in *EXTL2* KO cell models (Fig. 2B). Overall, these results suggest the existence of a complex regulatory interplay between CS and HS biosynthesis, supporting the hypothesis that the balance between HS and CS enzymatic machinery determines cellular HS/CS ratio (364). Interestingly, HS glycosyltransferases that act in the initial steps of the chain elongation, like EXTL3, appear to have a preponderant activity in this balance.

Additionally, we addressed the effects of altered GAGosylation over tumor cells motility and signaling events. Cellular functional analysis showed that *EXTL2* KO and concomitant HS increase and SDC4 overexpression promoted a more motile and invasive phenotype in MKN74 gastric cells (Fig. 4C-E). Previous reports have depicted the functional role of HS and HSPGs in modulating adhesion and migration in different cell lines (254, 366). SDC4 for example, displays important roles in motility as a key component of cells' focal adhesions. In line with our results, Gopal *et al*. revealed a direct correlation between HS chains and the role of SDC4 in mechanotransduction, showing that the expression of SDC4 modified with multiple HS chains is required for correct organization of cells' cytoskeleton actin components and formation of focal adhesions (362).

GPCs have also been previously associated to changes in tumor cell migration and invasion in various types of cancer (367, 368). GAG chains, HS in particular, were shown to be significantly important to the roles displayed by GPCs in cancer progression (343). Our data revealed that only *GPC1*, *GPC4* and *GPC5* were expressed in MKN74 gastric cancer cells. *GPC4* was downregulated in both KO models (Fig. 2A), therefore it is unlikely that it plays a role in promoting a more aggressive phenotype. In addition, no changes were observed on *GPC1* and *GPC5* upon KO of *EXTL2* (Fig. 2A). Still, these transcriptomic data do not allow to exclude that altered GPCs glycosylation features might contribute to the observed phenotype.

Notably, we observed that *EXTL3* KO cells presented unaltered migration and invasion (Fig. 4C-E). Since functional overlap between CS and HS has been previously described in different cell events, we speculate that increased CS content may compensate the absence of HS and re-establish *EXTL3* KO cell motility capabilities (369). Further studies are warranted to understand if CS acts on similar HS-governed signaling pathways or activates distinct signaling cascades that ultimately lead to similar functional behavior.

HS binds with high affinity to several biologically active proteins and contributes to the role of co-receptor displayed by many HSPGs by modulating ligand interactions with their targeted cell surface receptor. Particularly, HS can tether RTK ligands and promote the receptors activation and downstream signaling pathways (370, 371). Therefore, we have investigated the role of the specific HS profiles displayed by *EXTL2* KO and *EXTL3* KO cells in RTKs' activation. Interestingly, our results showed that abrogation of EXTL2 and subsequent cellular HS increase resulted in decreased EphA 4 activation in gastric cancer cells (Fig. 5). EphA 4 belongs to a large and unique family of Erythropoietin-producing hepatocellular (Eph) receptors whose members have been shown to modulate cell morphology, adhesion, migration, and invasion (372). EphA 4 activation has been previously associated with increased migration of cancer cells (373, 374), and particularly in gastric cancer, EphA 4 upregulation has been reported as a bad prognostic factor (375-377). However, the *EXTL2* KO model that displays a high motility phenotype shows decreased phosphorylation of this receptor (Fig. 5). The analysis of the interaction of heparin and HS with ephrin ligands and Eph receptors showed that only ephrin-A3 ligands bind to both, while EphA 4 receptor, which was only tested for heparin interactions, exhibited no detectable direct binding. In addition, it was reported that ephrinA3-HS binding was essential to mediate EphA 4 activation (378). HS sulfation is known to determine HS binding affinity and impact HS biological roles (379, 380). Therefore, we cannot exclude that the changes in HS sulfation detected in *EXTL2* KO (Fig. 3B) might contribute to impair HS-ligand and/or HS-receptor interactions and interfere with activation and that an alternative pathway might underlie the *EXTL2* KO motility features. Additional analyses are warranted to further understand the molecular determinants underlying HS-mediated EphA 4 activation and downstream regulation.

As illustrated in Fig. 6, our data revealed that lack of EXTL2 enzymatic activity results in the remodeling of GAGs on cell glycocalyx, promotes SDC4 upregulation and overproduction of HS with altered sulfation profile. Ultimately, these glycosylation changes were shown to promote a highly motile and invasive phenotype and to downregulate the activation of EphA 4 cell surface receptor. On the other hand, *EXTL3* KO blocks HS synthesis and triggers CS biosynthesis pathways. Our results showed that the absence of EXTL3 catalytic activity leads to the expression of SDC4 molecules abnormally modified with CS chains, which might compensate for the lack of HS and rescue cell functional behavior.

In conclusion, we showed that the abrogation of the tumor suppressor gene *EXTL2* in gastric cancer cells contributes to aggressive cellular features, such as increased motility and invasion. This further supports the importance of the molecular mechanisms underlying dysregulation of HS biosynthetic and postsynthetic modification pathways in cancer. Currently, there is very limited knowledge regarding the expression and impact of EXTL2 and EXTL3 in the context of cancer pathologies (284, 381, 382). Evaluating the expression profiles of these enzymes in clinical samples of gastric cancer patients at different stages of the disease could be of high clinical value as prognostic factors for patient's stratification or novel targets for therapy.

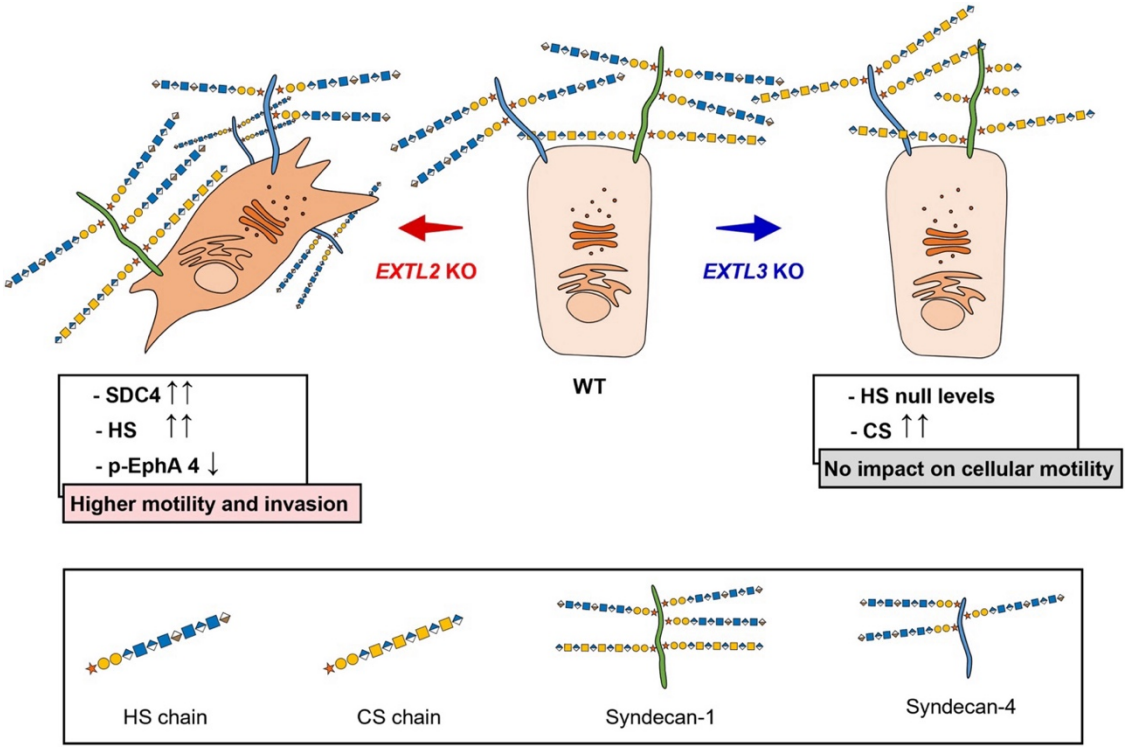


Figure 6. Schematic representation depicting the impact of EXTL2 or EXTL3 abrogation on gastric cancer cells' glycocalyx remodeling and functional features. GAG biosynthesis is regulated by EXTL2 and EXTL3. KO of *EXTL2* promoted HS overproduction and SDC4 overexpression in gastric cancer cells, ultimately inducing increased migration and invasion capabilities concomitant with lower EphA 4 activation. On the other hand, EXTL3 elimination triggered CS biosynthesis, while blocking HS biosynthetic pathways, and induced the expression of abnormal SDC4 molecules modified with CS chains, instead of HS, which might contribute to rescue cell motility properties. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

EXPERIMENTAL PROCEDURES

Cell lines, cell culture and genetic engineering

In this study, we resorted to the human gastric adenocarcinoma cell line MKN74, obtained from the Japanese Collection of Research Bioresources Cell Bank. Cells were cultured in RPMI–1640 (Gibco) culture medium, supplemented with 10% (v/v) Fetal Bovine Serum (FBS) (Biowest), at 37°C, under 5% (v/v) CO_2 conditions.

The glycoengineered *EXTL2* KO and *EXTL3* KO cell models were generated from the gastric cancer cell line MKN74, *via* CRISPR-Cas 9 genome editing. Briefly, cells were co-transfected with Cas9 endonuclease vector containing GFP and a plasmid with a validated gRNA for the target gene (Table 1). KO clones were obtained by single cell sorting of FACS enriched nuclease-expressing cells and gene KO was validated by indel detection using amplicon and RFLP combination analysis, as previously described (383). Two clones were selected for each target gene and gene KO confirmed by Sanger sequencing and indels were validated through Tracking of Indels by Decomposition (TIDE) methodology (Supplementary Fig. S1A and B) (384).

Table 1. *List of validated gRNAs used to target EXTL2 and EXTL3 based on (385) and results of indel analysis.*

Gene	gRNA sequence	Target exon	Indels	
EXTL2	CAGCTACCAGTAATAATACG	1/4	clone 3	clone 4
			+1	+1
EXTL3	GGTGGGGAACGAGCTGTGCG	1/5	clone 1	clone 5
			+1	+1

Antibodies

The antibodies and conditions used for the analysis described below are listed in Table 2.

Table 2*. List of antibodies and conditions used for WB, Immunofluorescence and Flow Cytometry (FACS) assays.*

Antibodies	**Applications and Dilutions**			**Source/Reference**
	WB	IF	FACS	
HS (F58-10E4)	1:600	1:200	1:600	Amsbio
HS stub region (F69-3G10)	1:500	1:100		Amsbio
SDC1 (B-A38)	1:125			Abcam
SDC4 (8G3)	1:2000			(386)
EphA 4 (4C8H5)	1:250			ThermoFisher
pEphA 4 [Tyr602] (EP2731)	1:1000			bioNova científica s.l.
EGFR (D38B1)	1:1000			Cell Signalling Technology
pEGFR [Tyr1068] (D75A)	1:1000			Cell Signalling Technology
pEGFR [Tyr1086] (36-9700)	1:2000			Invitrogen
α-Tubulin (DM1A)	1:10000			Sigma-Aldrich
Horseradish Peroxidase-conjugated α mouse IgG	1:5000			Jackson ImmunoResearch Laboratories, Inc
Horseradish Peroxidase-conjugated α mouse IgG1	1:8000			Jackson ImmunoResearch Laboratories, Inc
Horseradish Peroxidase-conjugated α mouse IgM	1:5000			Jackson ImmunoResearch Laboratories, Inc
Rabbit α-mouse Ig – FITC		1:70		DAKO
Alexa Fluor® 647 α mouse IgM			1:400	Jackson ImmunoResearch Laboratories, Inc

HS detection by Flow Cytometry Assay

Cells previously cultured in 75cm^2 flasks, were detached with versene (Gibco), resuspended in fresh medium and centrifuged at 300g for 5 min. Cells (1×10^6 cells/mL) were washed 2 times in Phosphate Buffered Saline (PBS), 1% Bovine Serum Albumin (BSA) (Sigma-Aldrich) solution, each wash followed by centrifugation at 300g for 5 min. The cell pellet was resuspended and incubated 1h at room temperature with the primary antibody 10E4 (Table 2) in PBS, and then washed again with PBS 1% BSA solution and incubated with APC fluorochrome-conjugate anti-mouse IgM. For negative controls, cell pellets were incubated with APC fluorochrome-conjugate anti-mouse IgM only. Lastly, cells were washed again and resuspended in PBS 1% BSA solution and filtered for analysis. Data was acquired using FACscanCanto II and analyzed with the FlowJo Software (v10). Mean Intensity of Fluorescence (MIF) values measured for the KO clones were normalized to the MIF values of WT, which were defined as a unit value. Three independent experiments were analyzed.

Fluorescence Microscopy

70-80% confluent cells seeded and grown on 12-well microscope glass slides (IBIDI) were washed twice with PBS and fixed on ice for 10 min with methanol (Thermo Fisher Scientific). Cells were washed again with PBS, blocked with Rabbit Serum (Dako), diluted in a ratio of 1:5 in PBS 10% BSA, and incubated with the primary antibody 10E4 (Table 2) in PBS 5% BSA, overnight at 4°C. This was followed by incubation with rabbit α mouse Ig – FITC conjugated secondary antibody for 30 min, at room temperature. Finally, cells were incubated with 100 µg/mL DAPI (4′, 6-diamidino-2-phenylindole), diluted 1:100 in PBS, and mounted in VectaShield mounting medium (Vector Laboratories). IF analysis for HS stubs was also performed on methanol fixed cells. After fixation, cells were permeabilized by incubation with 0.5% Triton X-100 in PBS for 20 min. Cells were washed with PBS and then incubated with 0.2M NH_4Cl in PBS for 20 min. For HS digestion cells were treated with Heparinase III (EC 4.2.2.8; 5 mU/mL) (Amsbio) in Tris-Buffered Saline (TBS) 0.1% BSA, supplemented with Calcium Acetate (0.1 mM), both diluted in 1:200 in water, for 2h at 37^{o}C. Cells were washed again with PBS, blocked with Rabbit Serum (Dako), diluted in a ratio of 1:5 in PBS 10% BSA, and incubated with the primary antibody 3G10 (Table 2) in PBS 5% BSA, overnight at 4°C. The following labelling steps were performed as previously described for 10E4 staining. Microscope images were obtained with the Zeiss Axio Imager Z1, Axiocam MR ver3.0 and Axiovision 4.8 Software (Carl Zeiss, Germany).

Western Blotting

Total protein lysates were obtained from confluent cells scraped and lysed with lysis buffer 17 (R&D Systems) supplemented with 1 mM sodium orthovanadate (Sigma-Aldrich), 1 mM phenylmethanesulfonyl fluoride (Sigma-Aldrich) and cOmplete™ protease inhibitor cocktail (Roche). The protein concentration of these lysates was determined using the DC protein assay (BioRad). Total protein lysates (30 µg for HS, 12,5 µg for HS stub, 25 µg for SDC1, SCD4 and EGFR, 35 µg for EphA 4 and 20 µg for p-EphA 4) were incubated with 4x Laemmli Sample Buffer (BioRad), supplemented with 10% β-mercaptoethanol, at 98ºC for 5 min. The protein samples were loaded and run in 4-15% precast polyacrylamide gels (BioRad) and blotted to a nitrocellulose membrane (GE Healthcare Life Sciences). The membranes were blocked for 1h at room temperature, with TBS 5% BSA and 0.05% Tween® 20 (Sigma-Aldrich) (TBS-T 0.05%) buffer, and then incubated overnight at 4°C with primary antibodies (Table 2) diluted in the blocking solution. Membranes were washed with TBS-T 0.05% and incubated for 1h at room temperature with horseradish peroxidase-conjugated secondary antibodies diluted in the blocking solution. New washes with TBS-T 0.05% were performed, membranes were developed, and the protein bands visualized with ECL chemiluminescent western blotting detection reagent and films (GE Healthcare Life Sciences). For WB analysis, immunoblot band densities related to the expression and phosphorylation of EphA 4 and EGFR receptors in each KO clone were normalized to the band density values of the WT. Receptor's phosphorylation levels were normalized to the levels of the total receptor. Results are represented as the mean values of the normalized phosphorylated receptor + SD. Three independent experiments were analyzed.

Glycosaminoglycan Enzymatic Digestion Assay

HS enzymatic digestion was performed by incubating 250 µg of protein lysates of each sample with Heparinase III (EC 4.2.2.8; 5 mU/mL) (Amsbio) in TBS 0.1% BSA, supplemented with Calcium Acetate (0.1 mM), overnight, at 37ºC with continuous mixing. 25µg of the resulting digested samples were then analyzed through WB analysis, by labelling HS digested resulting stubs with 3G10 antibody (Table 2). CS enzymatic digestion was performed by incubating 50 µg of protein lysates of each sample with Chondroitinase ABC (EC 4.2.2.4; 10 mU/µl) (Amsbio) in TBS 0.1% BSA, supplemented with 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$, overnight, at 37ºC with continuous mixing.

Real-time PCR analysis

For mRNA quantification/gene expression analysis, total RNA was extracted from MKN74 WT and glycoengineered cell models (*EXTL2* KO and *EXTL3* KO) using TRIzol Reagent (ABP Biosciences). 5 µg of RNA from each cell condition was converted into cDNA

via reverse transcription by incubation with random hexamers/primers, using the SuperScript™ IV Reverse Transcriptase Kit (Invitrogen). Real-time PCR was performed using for each condition, per well: 4 µL of cDNA diluted 1:10 in ultrapure water, 0.6 µL of each primer at 10 µM (Table 3), 10 µL of PowerUp™ SYBR™ Green Master Mix (Thermo Fischer Scientific) and ultrapure water up to a final volume of 20 µL. The real-time PCR run was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems). RQ values were determined for each gene and 18S ribosomal RNA was used as a housekeeping gene to normalize relative gene expression ($RQ = 2^{-(\text{target Ct} - \text{18S Ct})}$). Two independent experiments with triplicates were analyzed.

Table 3. *List of primers used in the real-time PCR analysis.*

Gene	Primer Fw	Primer Rv
SDC1	5'-ATGGCTCTGGGGATGACTCT-3'	5'-GTGGGAATAGCCGTCAGGAG-3'
SDC2	5'-AGGATGTAGAGAGTCCAGAGCT-3'	5'-TGTATCCTCTTCGGCTGGGT-3'
SDC3	5'-TGACATCCCTGAGAGGAGCA-3'	5'-GCTACCACCTCATTGGCTGT-3'
SDC4	5'-CCGGAGCCCTACCAGACGAT-3'	5'-AGGCACCAAGGGATGGACAA-3'
GPC1	5'-CATCGGGTGTGGAGAGTG-3'	5'-TGAGCGTGTCCCTGTTGTC-3'
GPC2	5'-CTGGGACACGACCTGGAC-3'	5'-GCCATCCAGTCATCTGCATAC-3'
GPC3	5'-CTGCTTCAGTCTGCAAGTATGG-3'	5'-GTGGAGTCAGGCTTGGGTAG -3'
GPC4	5'-AGTGTGGTCAGCGAACAGTG-3'	5'-CAAACATATCATTCAGGGATTTCTC-3'
GPC5	5'-GCCGCCCTGTAAGAACAC-3'	5'-TCATTCCATGCTTCTCTTTGC-3'
GPC6	5'-CCAGGCATAAGAAATTTGACG-3'	5'-CATGTACAGCATGCCATAGGTC-3'
18 S	5′-CGCCGCTAGAGGTGAAATTC-3′	5′-CATTCTTGGCAAATGCTTTCG-3′

GAG purification and HS/CS disaccharide analysis

To evaluate HS disaccharide composition, GAGs were isolated and purified, and structural analysis was performed as described in (387). For preparation and purification of GAG chains, cells were exhaustively digested with trypsin (45 min at 37°C). Cellular debris were discarded by centrifugation and supernatants were recovered and incubated with 250 U/µL Benzonase (Merck) and 2 mM $MgCl_2$ for 2 h at 37°C. Benzonase was then inactivated by heating supernatants at 96°C for 5 min and additional centrifugation was performed to discard nucleic acids. The recovered supernatants were applied to a DEAE-Sephacel column (Pharmacia Biotech) equilibrated in 20 mM phosphate pH 6.5. The column was then extensively washed with 20 mM phosphate pH 6.5, 0.3 M NaCl, then GAGs were eluted with 20 mM phosphate pH 6.5, 1 M NaCl. Samples were desalted over a Pd-10 column (GE Healthcare), lyophilized, and stored at -20°C prior to analysis. For HS disaccharide analysis, samples were resuspended in 100 mM sodium acetate, 0.5 mM calcium acetate, pH 7.1, then digested into disaccharides by incubation with a mix of heparinase I, II and III (10 mU each) for 48 h at 37°C. Compositional analysis were performed by RPIP-HPLC, by applying samples to a C18 reversed phase column equilibrated in a H_2O/acetonitrile (8.5%) buffer supplemented with 1.2 mM of ion-pairing tetra-*N*-butylammonium hydrogen sulfate (TBA),

then resolved using a multi-step NaCl gradient calibrated with HS disaccharide standards. On-line post-column disaccharide derivatization was achieved by the addition of 2-cyanoacetamide (0.25%) in NaOH (0.5%), followed by fluorescence detection (excitation 346 nm, emission 410 nm). For CS disaccharide analysis, purified samples were digested into disaccharides by incubation with 500 mU of Chondroitinase ABC in 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$, for 24 h at 37°C. The following compositional analysis was performed similarly to what was previously described for HS disaccharide analysis, using CS disaccharide standards instead.

Annexin V Viability Assay

To evaluate cell viability, $1.5x10^5$ cells per well were seeded on 6 well plates (Corning Incorporated Costar). After 48h in culture, cells were trypsinized (Biowest), resuspended on the medium they were cultured in and centrifuged at 300g for 5 min. The supernatant was discarded and the cell pellet was resuspended again in fresh medium. Cells were washed twice in PBS, followed by another two washes with Annexin V Binding Buffer (BioLegend). The resulting cell pellet was resuspended and incubated with Annexin V-FITC (BioLegend) diluted with a 1:40 ratio in Annexin V Binding Buffer for 15 min at room temperature. Cells were filtered and data was acquired using a BD FACSCanto II (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) and analyzed with the FlowJo Software (v10). Two independent biological replicates were analyzed.

Proliferation Assay

Cell proliferation was determined using Click-It™ Plus EdU Alexa Fluor™ 647 Flow Cytometry Assay Kit (Molecular Probes, Invitrogen, OR, USA) using the BD FACSCanto II flow cytometer. Cells were counted and adjusted to $2.5x10^5$ cells/mL and seeded in a $25cm^2$ flask in RPMI supplemented with 10% FBS and left to grow at 37ºC in 5% CO_2 atmosphere conditions. In parallel, cells were also grown in simple media (without FBS). 48h later, cells were incubated with 10 µM of EdU for 1h30 prior to harvesting. EdU non-treated cells served as control, while cells in simple media served as a cell arrest control. Detection of EdU incorporation into DNA was performed according to the manufacturer's instructions. Briefly, harvested cells (as described earlier) were washed in PBS 1% BSA and fixed in 100 µl Click-It1 fixative, for 15 min, at room temperature in the dark. Cells were then washed again in PBS 1% BSA and resuspended in 100 µL saponin-based permeabilization and wash reagent. Click-It1 EdU reaction cocktail was prepared according to the manufacturer's instructions and added. Samples were incubated for 30 min at room temperature in the dark and washed with saponin-based permeabilization and wash reagent. Cell pellet was resuspended in 300 µL saponin-based permeabilization and wash reagent, acquired using

BD FACSCanto II cytometer, and analyzed using FlowJo software (v10). Two independent biological replicates were analyzed.

Wound Healing Assay

$5.3x10^5$ cells/mL were seeded in each side of a silicon insert, previously adhered to a well of a µ-Slide 8 Well Collagen IV or Fibronectin coated (IBIDI). Cells were kept in culture for 24h, at 37◦C in 5% CO_2 atmosphere conditions. The inserts were then removed, cells were washed with RPMI–1640 culture medium supplemented with 10% FBS, and fresh supplemented medium was added to each well. Time-lapse microscopy was performed using Leica DMI6000 (Wetzlar) and three bright field images per well/condition were acquired for 24h with intervals of 10 min. The wound healing rate was evaluated by measuring the total wound area at each time point using the ImageJ software. Migration assay results were depicted as the average values of the % of closing wound + SD. The % of closing wound was calculated by subtracting the area of the open wound at the first time point (t=0h) to the area determined for each time point, followed by normalization of the resulting values to the wound area determined for the first time point (t=0). Two independent biological experiments with at least technical triplicates were analyzed.

Matrigel Invasion Assay

Invasion assays were performed resorting to a 24-well plate of BD BioCoat MatrigelTM Invasion Chambers (BD Biosciences). $2x10^5$ cells were initially seeded and incubated in RPMI–1640 medium in the upper chamber for 24h at 37°C in 5% CO_2 atmospheric conditions. The lower side of the well contained only RPMI– 1640 medium supplemented with 10% FBS. The Matrigel coated chambers were then washed with PBS, the non-invasive cells adhered on the inner side of the chamber were removed with a cotton swab, and the chambers were fixed in ice-cold methanol for 10 min and air dried. The chambers were washed with PBS, the Matrigel coated membranes were removed and mounted in glass coverslips with VectaShield mounting medium with DAPI (Vector Laboratories). Microscope images of the stained nuclei were obtained resorting to the Zeiss Axio Imager Z1, Axiocam MR ver3.0 and Axiovision 4.8 Software (Carl Zeiss, Germany) and the total number of invasive nuclei was counted. Invasion assay results were represented as the average values of the fold changes of the number of invasive cells + SD. The number of invasive cells of the KO models was normalized to WT, which was defined as a unit value. Two independent biological experiments with technical triplicates were analyzed.

Receptor Tyrosine Kinase (RTK) phosphorylation array

The activation state of important RTKs was determined by using the Proteome Profiler Human Phospho-RTK Array Kits (R&D Systems #ARY001B) following the manufacturer's instructions. Briefly, confluent cells ($1x10^7$ cells/mL) were solubilized in lysis buffer 17 (R&D Systems) supplemented with cOmpleteTM protease inhibitor cocktail (Roche) and PhosSTOP phosphatase inhibitor (Sigma-Aldrich). The protein concentration of these lysates was determined using the DC protein assay (BioRad). The RTK array membranes were incubated with the Array Buffer 1 for 1h at room temperature with continuous shacking. 300 µg of whole cell lysates were diluted in Array buffer 1 to a final volume of 1.5 mL and incubated on the RTK array membranes overnight at 4°C with continuous shaking. Membranes were then washed thrice with 1x Wash Buffer, 10 min each wash, followed by incubation with 1:5000 Anti-Phospho Tyrosine-HRP detection antibody diluted in Array buffer 2, for 2h at room temperature with continuous shaking. Membranes were washed again thrice with 1X Wash Buffer and finally were developed and visualized with ECL chemiluminescent detection reagent and films. For dot quantification, densitometry was evaluated and the phosphorylation fold change of each RTK was calculated for each KO cell model by comparison with the RTK phosphorylation values determined for the WT. The following calculation was performed: Fold change = 1 – (average WT p-RTK densitometry/ average KO p-RTK densitometry).

Statistical Analysis

Statistical analysis was carried out using GraphPad Prism 6 and statistical significance was considered when P values were ≤0.05 (* means p≤0.05; ** means p≤0.01; *** means p≤0.001; **** means p≤0.0001). For flow cytometry, invasion assay and WB analysis, statistical significance of WT versus *EXTL2* KO and WT versus *EXTL3* KO was determined using unpaired Student's t-test with Welch's correction, with a 95% interval of confidence. For real-time PCR statistical significance was determined with one-way ANOVA using Tukey's test for multiple comparisons. Migration assay statistical significance was calculated by two-way ANOVA with a 95% interval of confidence and shown for each KO cell model relative to the WT.

DATA AVAILABILITY

All data described and discussed are contained within the manuscript.

SUPPORTING INFORMATION

This manuscript contains supporting information.

ACKNOWLEDGMENTS

The authors would like to acknowledge Guido David for the 8G3 antibody and Henrik Clausen for providing the gRNAs plasmids for *EXTL2* and *EXTL3* KO and GFP-Cas9. The authors acknowledge the support of the i3S Scientific Platforms Advanced Light Microscopy, member of the PPBI (PPBI-POCI-01-0145-FEDER-022122), and Translation Cytometry.

AUTHOR CONTRIBUTIONS

C.M. and A.M. conceptualization; C.M., J.P. and I.F.R. data acquisition and analysis; C.G. CRISPR-Cas9 KO models development; R.R.V. structural analysis and supervision; C.M. and A.M. writing original draft, J.P., C.G., I.F.R., R.R.V. and C.A.R. manuscript writing-review; C.G., C.A.R., R.R.V. and A.M. resources and funding acquisition. All authors read and approved the final manuscript.

FUNDING AND ADDITIONAL INFORMATION

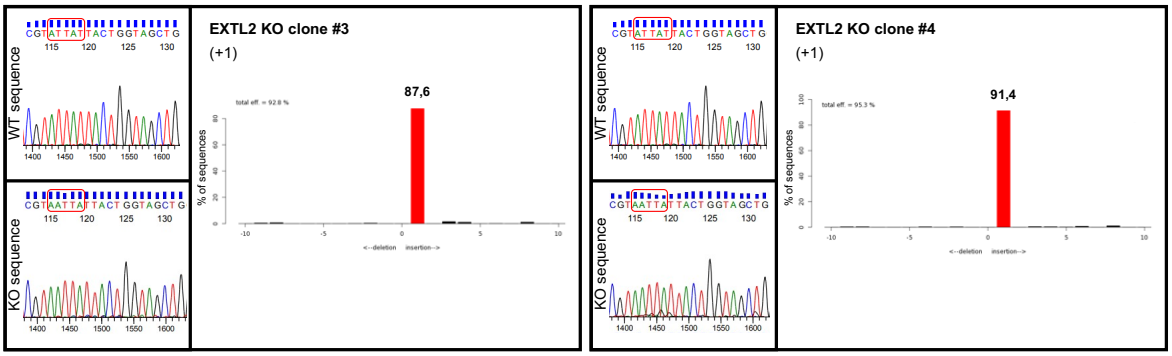
This work was funded by FEDER funds through the Operational Programme for Competitiveness Factors-COMPETE (POCI-01-0145-FEDER-007274, POCI-01-0145-FEDER-028489 and POCI-01-0145-FEDER-029780) and National Funds through the Foundation for Science and Technology (FCT), under the projects: PTDC/MED-ONC/28489/2017 (to AM), PTDC/MED-QUI/29780/2017 (to CG), UID/BIM/4293 and the project NORTE-01-0145-FEDER-000051, supported by Norte Portugal Operational Regional Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). COST Action CA18103 INNOGLY: INNOvation with Glycans new frontiers from synthesis to new biological targets (ECOST-STSM-Request-CA18103-45514). CM and JP acknowledge FCT PhD scholarships 2020.06412.BD and SFRH/BD/137319/2018, respectively. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). This work was also supported by the CNRS and the GDR GAG (GDR 3739), the “Investissements d’avenir” program Glyco@Alps (ANR-15-IDEX-02), by a grant from the Agence Nationale de la Recherche (ANR-17-CE11-0040). IBS acknowledges integration into the ©nterdisciplinary Research Institute of Grenoble (IRIG, CEA).

CONFLICT OF INTEREST

The authors declare no competing financial interests.

SUPPORTING INFORMATION

A



B

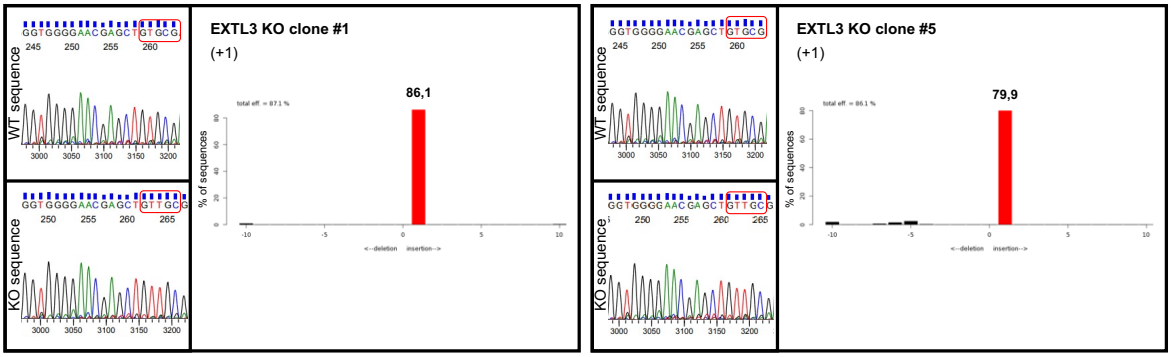
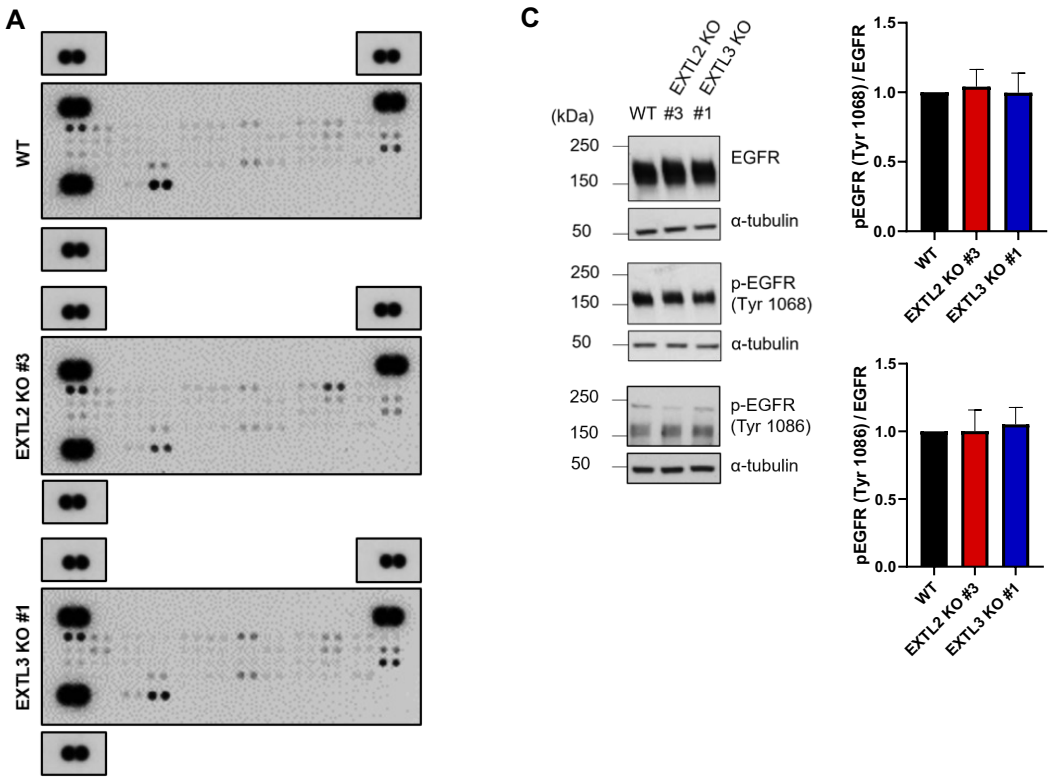


Figure S1. Validation of the glycoengineered *EXTL2* KO and *EXTL3* KO cell models generated *via* CRISPR-Cas 9 genome editing. KO validation through Sanger sequencing for **(A)** *EXTL2* KO clones #3 and #4 and **(B)** *EXTL3* KO clones #1 and #5. Indel sequencing was validated through Tracking of Indels by Decomposition (TIDE) methodology. The MKN74 WT sequence (top left panel) was used as the reference (control) for comparison with all KO clones.



B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	Reference spot																						Reference spot	
B	EGFR		ErbB2		ErbB3		ErbB4		FGF R1		FGF R2α		FGF R3		FGF R4		Insulin R		IGF-I R		Axl		Dtk	
C	Mer		HGF R		MSP R		PDFG Rα		PDFG Rβ		SCF R		Flt-3		M-CSF R		c-Ret		ROR1		ROR 2		Tie-1	
D	Tie-2		TrkA		TrkB		TrkC		VEGF R1		VEGF R2		VEGF R3		MuSK		EphA 1		EphA 2		EphA 3		EphA 4	
E	EphA 6		EphA 7		EphB 1		EphB 2		EphB 4		EphB 6		ALK		DDR 1		DDR 2		EphA 5		EphA 10			
F	Reference spot				EphB 3		RYK																Control (-) PBS	

Figure S2. Assessment of RTKs phosphorylation state in MKN74 *EXTL2* KO and *EXTL3* KO cancer cell models. (A) Evaluation of the phosphorylation state of 49 RTKs expressed in MKN74 WT, *EXTL2* KO and *EXTL3* KO cells. Each couple of dots represents the staining intensity for each p-RTK. Less exposed images of the reference spots are shown above and below each blot. **(B)** Schematic illustration showing the location of each RTK specific antibody in the human p-RTK array. **(C)** The activation of EGFR was measured by WB analysis using antibodies that specifically recognize EGFR Tyr-1068 and Tyr1086. α-Tubulin was used as loading control. Band density values related to EGFR expression and phosphorylation were normalized to the band density values of the WT, which were defined as one unit. p-EGFR levels were normalized to the levels of the total receptor. Bar graph shows the mean values of the normalized phosphorylated receptor + SD. n=3 independent biological assays.

Additional results

Addressing the role of Syndecan-4 in the glycosaminoglycan profile exhibited by the *EXTL2* or *EXTL3* KO glycoengineered gastric cancer cell models

Syndecan-4 is a main player in determining MKN74 Heparan Sulfate and Chondroitin Sulfate levels

INTRODUCTION

Upon understanding the impact that HS glycosyltransferases EXTL2 and EXTL3 have on gastric cancer cells' GAG profile and, consequently, on cellular motility (388), we aimed to further comprehend the molecular mechanism driving this tumor cell aggressive behavior. From the work described in the present chapter, we have concluded that the KO of *EXTL2* in MKN74 gastric tumor cells induced increased levels of cellular HS, with decreased 6-*O*-sulfation motifs, and overexpression of *SDC4*. However, it remained unclear whether there was a specific main driver of the observed aggressive phenotype or if the cells' general glycoprofile ultimately influenced cell motility.

In a recent work by our group, Poças *et al.* have determined that SDC4 is upregulated in gastric tumors, from the intestinal subtype, and it associates with patients' poor survival. In addition, and resorting to MKN74 *SDC4* KO glycoengineered cells, we showed that SDC4 promoted cell motility and invasion capabilities, and it defined extracellular vesicles' tropism to common metastatic organs of gastric cancer (270). Having these data in consideration, we aimed to evaluate the specific influence of SDC4 in the GAG profile exhibited by the *EXTL2* or *EXTL3* KO glycoengineered gastric cancer cell models and the impact on the cell phenotype.

RESULTS

SDC4 is a major HS carrier in gastric cancer cells

To address the proposed aim, we established double KO models by abrogating SDC4 from the previously established MKN74 *EXTL2* KO (*EXTL2* and *SDC4* dKO) and *EXTL3* KO (*EXTL3* and *SDC4* dKO) cells, *via* the CRISPR-Cas9 methodology. SDC4 protein levels were determined through WB analysis, confirming the successful KO of *SDC4* in these cells, as we could not detect neither the glycosylated nor non-glycosylated forms of this PG (Fig. 7A). Additionally, total levels of cellular HS were also determined *via* WB analysis. Noticeably, the repression of *SDC4* in *EXTL2* KO cells significantly reduced HS to lower levels than the WT cells, while no significant changes in HS labeling were detected in *EXTL3/SDC4* dKO cells, as expected (Fig. 7B). This result suggests that SDC4 is a major HS carrier in MKN74 gastric cancer cells. Mild HS staining was still present at higher molecular weights (> 250 kDa) in *EXTL2/SDC4* dKO cells, similar to WT, which was not observed in the *EXTL3/SDC4* dKO cell lysates. Therefore, we cannot disregard that other HS modified PGs might be present in this cell model, considering also additional voluminous HSPGs that might not be detected *via* WB analysis.

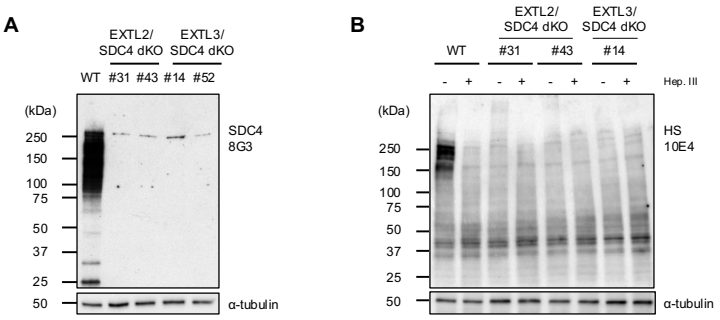


Figure 7. *EXTL2/SDC4* dKO gastric cancer cells display altered levels of HS. (A) To validate the established MKN74 *EXTL2/SDC4* dKO and *EXTL3/SDC4* dKO cell models, the protein levels of SDC4 were evaluated by WB analysis, using the 8G3 mAb. **(B)** HS total content in MKN74 WT, *EXTL2/SDC4* dKO and *EXTL3/SDC4* dKO cells was evaluated by WB analysis. Protein lysates were treated with or without Heparinase III (Hep. III) and immunoblotted for undigested HS using the 10E4 mAb**.** Images are representative of independent duplicates.

Overproduction of CS in MKN74 *EXTL3* KO cells is SDC4-independent

We previously observed that in the absence of *EXTL3* expression, leading to null HS production, MKN74 cells expressed SDC4 aberrantly modified with CS chains (Chapter 2, Fig. 2B). This was further supported by disaccharide structural analysis, showing a trend towards increased CS/DS levels in the *EXTL3* KO cells (Chapter 2, Fig. 3E). To further validate these results, and to learn about other possible new CS-carriers in the glycoengineered cell lines, we evaluated CS levels in the single and double KO models Protein lysates were digested with Chond. ABC and both non-digested and digested samples were immunoblotted against C4S and C6S using the BE-123 and MK-302 mAb, respectively (Fig. 8 A and B). While no significant changes were determined in C4S labeling upon KO of *EXTL2,* a significant increase was noticed in *EXTL3* KO cells, validating our previous observations. Moreover, several C4S signaling bands across the blot were significantly more intense in the Chond. ABC digested *EXTL3* KO protein lysates (Fig. 8 A). The most intense band showed up particularly close to 37 kDa which, considering the previous data (Chapter 2, Fig. 2B), is likely due to the presence of CS-modified SDC4. As we performed extensive Chond. ABC digestion of the protein lysates (0.5U/mL overnight, 37ºC), we assumed the labeling in the digested conditions was from very short CS stubs. Therefore, the labeling appearing at higher molecular weights in *EXTL3* KO cells, that lack HS, might indicate enrichment in PGs bearing other GAGs besides CS.

Regarding the labeling of C6S, the obtained results were quite surprising. Although changes in MK-302 immunolabeling were observed between non-digested and digested samples in all conditions (Fig. 8B), C6S disaccharides were not detected when performing disaccharide structural analyses (Chapter 2, Fig. 3E). This suggests that the labeling observed in this WB is likely unspecific and related to the processing of samples during Chond. ABC digestion.

Next, we analyzed C4S labeling in the dKO cell models. Interestingly, significantly increased labeling was detected in the Chond. ABC digested *EXTL3/SDC4* dKO protein lysates, although the lower band at 37 kDa was no longer detected, further confirming the absence of SDC4 in these cell models (Fig. 8C).

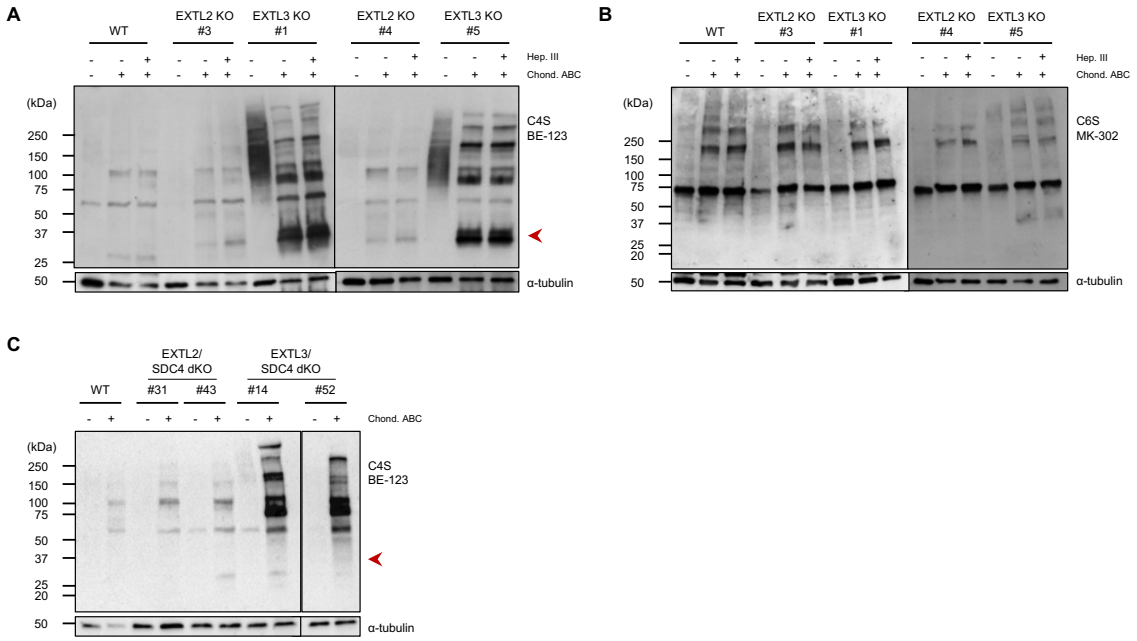


Figure 8. MKN74 *EXTL3* KO and *EXTL3/SDC4* dKO cells present increased levels of CS in different protein carriers. The glycoengineered *EXTL2* KO and *EXTL3* KO cell models, either **(A and B)** single KO or **(C)** double KO for *SDC4*, were evaluated for CS total content *via* WB analysis. Protein lysates were treated with or without Chondroitinase ABC (Chond. ABC) and Heparinase III (Hep. III), to enhance the antibody labeling, as advised by the antibodies manufacturer, and immunoblotted for Chondroitin-4-sulfate (C4S) and Chondroitin-6-sulfate (C6S) motifs using the **(A and C)** BE-123 mAb and **(B)** MK-302 mAb, respectively. **(A and C)** The position of CS modified SDC4 is indicated with red arrowheads.

DISCUSSION

SDC4 is an ubiquitous cell surface HSPG, exerting pivotal roles in homeostasis and disease (259). It is known to be a key component of focal adhesions, favoring the ECM-cell cross-talk and cytoskeleton remodulation in response to extracellular cues, in a HS-dependent manner, ultimately influencing cell motility and migration (362, 389).

Aligned with the previously work reported by our team associating SDC4 with gastric tumor aggressive behavior (270), we observed that *EXTL2* KO cells, harboring increased levels of HS but also increased expression of *SDC4*, presented a highly motile and invasive phenotype (388). This prompted us to try to understand the specific impact of SDC4 in the glycoengineered tumor cells' GAGosylation patterns that associate with this aggressive cell behavior.

Our results highlighted SDC4 as a major cell surface HS carrier in MKN74 cells, as *EXTL2/SDC4* dKO cells displayed drastically reduced HS labeling (Fig. 7B). Additionally, these data further suggest that EXTL2 negatively regulates the expression of this specific

HSPG and, as a direct consequence, influences overall cellular HS levels. Although EXTL2 is hypothesized to be a general negative suppressor of HS biosynthesis (148), our data suggests otherwise, revealing that lack of SDC4 overrides over the upregulation of HS observed in *EXTL2* KO cells (Fig. 7B). Nonetheless, due to the previously described abnormal expression of SDC4 in the context of gastric malignancies, we cannot exclude that the observed high "SDC4/all HSPGs" ratio in the MKN74 cell line (Chapter 2, Fig. 2A) might influence the activity and substrate specificity of EXTL2. It would be important to further evaluate this possible EXTL2/SDC4 regulatory loop in different cell lines. In addition, the influence of *SDC4* KO over the HS GAGosylation profile of MKN74 *EXTL2* KO cells does not allow to compare the effects of high HS (*EXTL2* KO) against the *EXTL2/SDC4* dKO.

Finally, our results confirmed that in the absence of HS biosynthesis, MKN74 cells GAGosylation profile undergoes a shift towards increased CS content (Fig. 8A). Moreover, CS evaluation in the *EXTL3/SDC4* dKO cell models showed that SDC4 might not be the only HSPG undergoing this HS to CS shift in the *EXTL3* KO cells. We also cannot rule out that other CSPGs might be overexpressed in these cells as a compensatory mechanism in response to the lack of HS biosynthesis.

Overall, our results highlight SDC4 as a major cell surface HS-carrier in gastric cancer cells and support the existence of a EXTL2/SDC4 regulatory loop that controls total cellular levels of HS. Moreover, we identified a potential cell dynamic GAGosylation profile involving different HSPGs, in addition to SDC4, that appear to undergo a shift from HS to CS in the absence of HS biosynthesis. These features should be addressed in the medical setting, as they might hold valuable clinical potential in cancer diagnosis and therapy.

EXPERIMENTAL PROCEDURES

Cell lines, cell culture, and genetic engineering

The human gastric adenocarcinoma cell line MKN74 was obtained from the Japanese Collection of Research Bioresources Cell Bank. Cells were cultured in RPMI1640 (Gibco) culture medium, supplemented with 10% (v/v) fetal bovine serum (FBS) (Biowest), at 37°C, under 5% (v/v) CO_2 conditions.

The glycoengineered double KO cells, *EXTL2/SDC4* dKO and *EXTL3/SDC4 d*KO cell models were generated from the MKN74 *EXTL2* KO and *EXTL3* KO gastric cancer cell models, respectively, *via* CRISPR-Cas 9 genome editing. Briefly, cells were co-transfected with Cas9 endonuclease vector containing GFP and a plasmid with a validated guide RNA for the target gene (5'-CGGAGCCCTACCAGACGATG-3'). KO clones were obtained by single cell sorting of fluorescence-activated single cell sorting enriched nuclease-expressing and validated as previously described (388).

Western Blotting

Total protein lysates were obtained from confluent cells scraped and lysed with lysis buffer 17 (R&D Systems) supplemented with 1 mM sodium orthovanadate (Sigma–Aldrich), 1 mM PMSF (Sigma–Aldrich), and cOmplete protease inhibitor cocktail (Roche). The lysates protein concentration was determined using the DC protein assay (Bio-Rad). Total protein lysates (30 µg for HS and 25 µg for SCD4, C4S and C6S) were incubated with 4× Laemmli sample buffer (Bio-Rad), supplemented with 10% β-mercaptoethanol, at 98°C for 5 min. The protein samples were loaded and run in 4% to 15% precast polyacrylamide gels (Bio-Rad) and blotted to a nitrocellulose membrane (GE Healthcare Life Sciences). The membranes were blocked for 1h at room temperature, with TBS 5% BSA with 0.05% Tween 20 (TBS-T 0.05%; Sigma–Aldrich) buffer, and then incubated overnight at 4°C with primary antibodies (Chapter 2, Table 2) diluted in the blocking solution. Membranes were washed with TBS-T 0.05% and incubated for 1h at room temperature with horseradish peroxidase–conjugated secondary antibodies diluted in the blocking solution. New washes with TBS-T 0.05% were performed, membranes were developed, and the protein bands visualized with ECL chemiluminescent WB detection reagent and films (GE Healthcare Life Sciences).

Glycosaminoglycan enzymatic digestion assay

HS enzymatic digestion was performed by incubating 250 µg of protein lysates of each sample with Heparinase III (EC 4.2.2.8; 5 mU/mL; Amsbio) in TBS 0.1% BSA, supplemented with Calcium Acetate (0.1 mM), overnight, at 37ºC with continuous mixing. CS enzymatic digestion was performed by incubating 50 µg of protein lysates of each sample with Chondroitinase ABC (EC 4.2.2.4; 10 mU/µl; Amsbio) in TBS 0.1% BSA, supplemented with 50 mM Tris–HCl pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$, overnight, at 37°C with continuous mixing.

Chapter 3

Addressing the role of Heparan Sulfate (HS) 6-*O*-sulfation in the regulation of HS and HS Proteoglycans biosynthesis and cellular distribution in gastric cancer

Disclosing the regulatory role of Heparan Sulfate 6-*O*-sulfation patterns in gastric cancer

Catarina Marques[1,2,3,4] Ana Raquel Ribeiro[1,2,5], Catarina Gomes[1,2], Rafaela Abrantes[1,2], Evelyne Gout[4], Bruno Cavadas[1], Celso Reis[1,2,5,6], Romain R. Vivès[4], Ana Magalhães[1,2,6]

1 - i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto 4200-135, Portugal.

2 - IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto 4200-465, Portugal.

3 - Programa Doutoral em Biologia Molecular e Celular (MCbiology), Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal.

4 - Univ. Grenoble Alpes, CNRS, CEA, IBS, Grenoble, France.

5 - FMUP - Faculdade de Medicina da Universidade do Porto, Porto 4200-319, Portugal.

6 - ICBAS - Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto 4050-313, Portugal.

Manuscript in preparation

ABSTRACT

Cell surface Heparan Sulfate (HS) Proteoglycans (HSPGs) are important players in various physiological and pathological events, by participating in cell-cell and cell-ECM crosstalk, influencing cell signaling cascades and controlling cell proliferation, adhesion, motility, and invasion. These key biological functions of HSPGs are known to be mainly dictated by HS structural features, namely sulfation patterns. Particularly in cancer, aberrant HS sulfation profiles affect tumor cell behavior in different contexts. Our work focuses on the regulation of HS biosynthesis, aiming to explore the role of 6-*O*-sulfation in fine-tuning HS and HSPGs expression and cellular/subcellular distribution in gastric cancer.

Here, we focused on two enzymes that modulate HS 6-*O*-sulfation and established 6-*O*-sulfotransferase *HS6ST1* knock-out (KO) and 6-*O*-endosulfatase *HSULF2* overexpression (OE) glycoengineered gastric cancer cell models. We demonstrated that these enzymes, by modulating HS 6-*O*-sulfation differently, promote the formation HS epitopes endowed with different biological functions. We revealed that ablation of intracellular HS 6-*O*-sulfation induced HS biosynthesis and overexpression of the cell surface HSPG Syndecan-4 (SDC4). In addition, we also showed that reduced HS 6-*O*-sulfation influences HSPG cellular localization. Finally, our results indicated that HSULF2 is highly expressed in gastric cancer and it associates with patients' worse prognosis, highlighting the clinical potential of this enzyme in gastric clinical management.

INTRODUCTION

Heparan Sulfate (HS) Proteoglycans (HSPGs) are important glycoconjugates expressed virtually in all mammalian cells. These macromolecules are particularly abundant in cell glycocalyx and extracellular matrix, and can also be transported in secreted extracellular vesicles (83-85). Depending on their localization, HSPGs can regulate various cell events, including the formation of extracellular gradients, cellular growth and proliferation, cell adhesion, migration and invasion, membrane trafficking and angiogenesis, playing a crucial role in physiology and pathology (86, 87).

HSPGs protein backbone is modified with long and linear HS polysaccharide chains composed of up to 200 repeating disaccharide units, each formed by *N*-acetylglucosamine (GlcNAc) attached to either glucuronic acid (GlcA) or iduronic acid (IdoA) residues (4). Although the core protein of HSPGs exhibits some intrinsic properties, most biological functions displayed by these glycoconjugates, including their participation in numerous signaling pathways, have been associated to HS-binding to various soluble molecules, including growth factors, chemokines, cytokines and morphogens, extracellular structural proteins, enzymes and transmembrane signaling receptors (94, 210). The binding affinity of HS to these proteins is influenced by their structure and high negative charge, conferred

by the presence of multiple sulfate groups, which results from a tightly fine-tuned process of biosynthesis.

HS biosynthesis occurs at the interface between endoplasmic reticulum and Golgi apparatus and in the Golgi apparatus, and it comprises three main stages: 1) formation of the tetrasaccharide linker, whereby HS chains are covalently attached to PG protein backbone; 2) HS initiation and elongation, comprising the addition of the first GlcNAc residue that orients the whole pathway towards synthesis of HS, followed by the polymerization of long non-modified HS chains; 3) extensive structural modifications, including epimerization and sulfation reactions, that promote the maturation of HS. In addition to the modifications that occur during biosynthesis, HS can be subjected to post-synthesis editing reactions that further change the sulfation patterns and length of mature chains (90). While the first two stages mainly involve the participation of multiple glycosyltransferases and lead to the polymerization of unmodified long chains, it is the third stage of modification reactions and the post synthesis editing enzymes that contribute to HS high structural variability and ultimately influence HS interactions and biological roles. Briefly, during polymerization, pro-heparan chains are targeted by *N*-Deacetylase/*N*-Sulfotransferases 1-4 (NDSTs 1-4) that catalyze the deacetylation and sulfation of GlcNAc residues. This step is followed by the epimerization of GlcA into IdoA by Glucuronyl C5-epimerase (GLCE), sulfation of IdoA/GlcA by 2-*O*-sulfotransferase 1 (HS2ST1), and sulfation of GlcNS/GlcNAc by 6-*O*-sulfotransferases 1-3 (HS6STs 1-3) and 3-*O*-sulfotransferases 1-6 (HS3STs 1-6) (135). The mature HS chains are then expressed at the cell membrane, while attached to their respective protein backbone, and can be targeted by extracellular 6-*O*-endosulfatases 1 and 2 (HSULFs 1-2), which catalyze the hydrolysis of HS 6-*O*-sulfate groups from highly sulfated disaccharides (IdoA2S-GlcNS6S) (313). Furthermore, Heparanase 1 (HPSE) can further process and depolymerize HS chains through the hydrolysis of internal GlcA–GlcNS linkages (176).

The biological relevance of HS sulfation patterns has been highlighted and studied in depth over the years, and considerable advances have been made to understand the regulation of HS sulfation motifs and their contribution to cell mechanisms related to health and disease. HS 6-*O*-sulfation has garnered a lot of interest, as it is the only modification that can be regulated both intracellularly, during biosynthesis, and extracellularly through post-synthesis mechanisms (163). In addition, this late modification classifies as a moderately abundant HS structural feature and may provide intermediate binding selectivity, as opposed to the more abundant and less selective *N*- and 2-*O*-sulfates and to the rarer and highly selective 3-*O*-sulfates (349). Fibroblast Growth Factor 1 (FGF1) and FGF2 are important examples of soluble ligands that rely on the binding to HS, and on the role that these polysaccharide chains play as a cell-surface co-receptors, to further promote

FGF-FGFR binding and activation of downstream signaling pathways. The presence of 6-*O*-sulfated residues is an important requirement for the establishment of the biologically active FGFR-FGF-HS ternary complex that promotes the activation of these signaling cascades (214, 218, 219, 221, 222). In accordance, it has been reported that HS 6-*O*-sulfation remodeling by HSULFs influences FGF2 signaling cascades (190, 390). Likewise, receptor-binding and signaling activation by several other growth factors (e.g. heparin-binding epidermal growth factor (HB-EGF), vascular endothelial growth factor (VEGF) and morphogens (e.g. bone morphogenic protein (BMP) and Wnt) are influenced by the presence of specific HS 6-*O*-sulfate residues (163, 349).

Noteworthy, significant changes in HS 6-*O*-sulfation degree and aberrant expression of enzymes involved in the regulation HS 6-*O*-sulfation have been reported in different pathological conditions. Particularly in the context of cancer, upregulation of *HS6ST1* and *HS6ST2* were associated with chondrosarcoma progression, together with higher HS 6-*O*-sulfation, determined in high-grade chondrosarcoma cell lines (297). Overexpression of *HS6ST2* and *HS6ST3* was also reported in colorectal (298) and breast (299) cancer, respectively, and higher HS 6-*O*-sulfation levels were found in the ovarian endothelium of patients with advanced stage cancer (300). Also in gastric cancer, increased expression of *HS6ST2* was associated with patients' poor prognosis (331). Similarly to the trend observed regarding the expression of 6-*O*-sulfotransferases, various reports have also described the upregulation *HSULF2* in several types of cancer, including lung (318), pancreatic (319), breast (320), colorectal (321), gastric (335), and liver carcinoma where its expression was associated with patients' worse prognosis (322). On the other hand, data have been more contradictory regarding HSULF1 expression. Unlike HSULF2, which is often associated with a pro-oncogenic role, HSULF1 expression and roles vary depending on the type of cancer, although it is more frequently described as having a tumor suppressor effect. *HSULF1* downregulation has been described in head and neck squamous cell carcinoma cell lines (314), hepatocellular carcinoma (315), breast and gastric cancer (333), amongst other. In opposition, and contradicting some of these data, increased expression of *HSULF1* was detected in pancreatic cancer cells (316), gastric cancer (335), leukemia and renal carcinoma (317). It is interesting to notice that HS6STs and HSULF2 are generally described as overexpressed in the context of different types of cancer. Taking in consideration the effects of these two enzymes in HS biosynthesis and 6-*O*-sulfation motifs - HS6STs promote overall HS 6-*O*-sulfation, while HSULFs are responsible for the removal of 6-*O*-sulfate residues within HS highly sulfated domains - this might suggest that an extremely fine HS/HSPGs regulation is carried that might influence tumorigenesis. Although there are many studies reporting the general expression of HS 6-*O*-sulfation enzymes in cancer and proposing signaling pathways that might be impacted by these features, the

specific function of HS 6-*O*-sulfation patterns in cancer and the regulatory mechanisms underlying these changes remain to be fully elucidated.

In our previous report, we showed that the KO of the glycosyltransferase Exostosin-Like 2 (EXTL2), that has been proposed to be a negative regulator of HS biosynthesis, induced overproduction of HS and overexpression of SDC4. Interestingly, HS disaccharide structural analyses further revealed that the newly produced HS had diminished 6-*O*-sulfated disaccharide levels. Overall, these changes were associated with pro-tumorigenic features, namely increased migration and invasion of gastric cancer cells (388).

Taking all of these data into consideration, this study aimed to clarify the biological significance of the enzymes that directly regulate HS 6-*O*-sulfation, namely HS6STs and HSULFs, and the general impact of HS 6-*O*-sulfation motifs in gastric cancer cell signaling and aggressive behavior.

RESULTS

Modulation of *HS6ST1* and *HSULF2* expression tunes HS 6-*O*-sulfation profiles

Initially, we screened a panel of six gastric cancer cell lines to learn about the expression levels of the genes coding for the 5 enzymes that regulate HS 6-*O*-sulfation: HS6STs 1-3 and HSULFs 1-2 (Fig.1A). We observed that only *HS6ST1* and *HSULF2* were expressed at significant levels across the six selected gastric cancer cell lines (Fig.1 B). Therefore, these enzymes were selected to study the impact of HS 6-*O*-sulfation in gastric cancer in the following experiments.

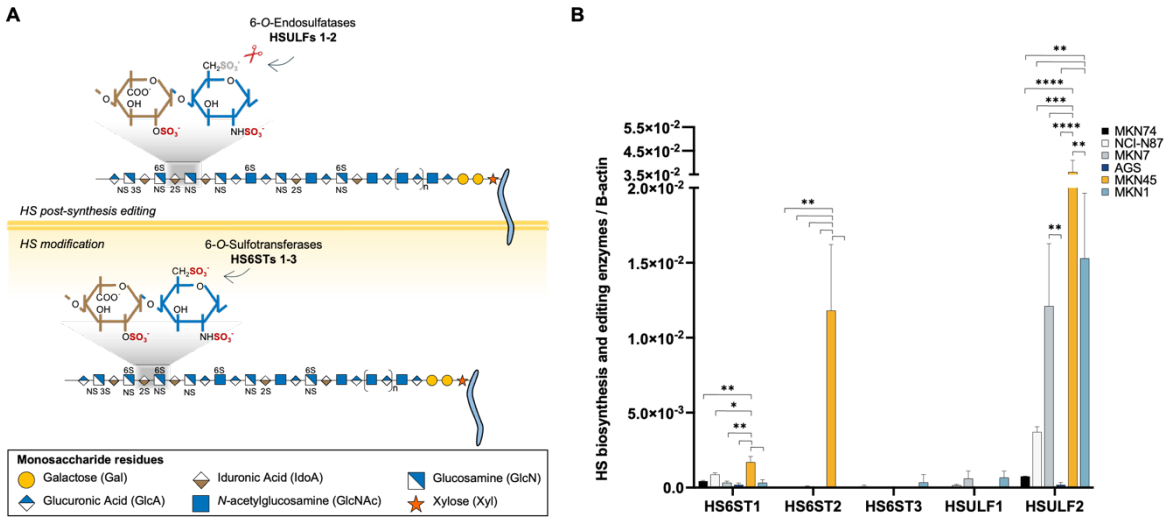
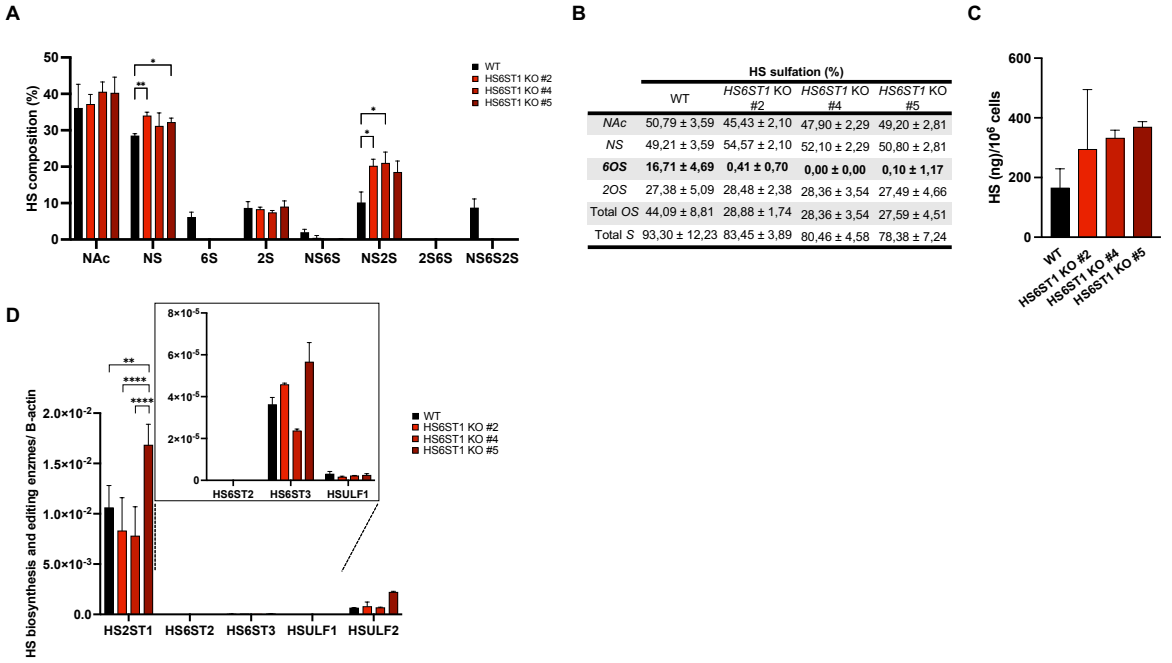


Figure 1. Transcriptomic analysis of the genes coding for HS biosynthesis and post-synthesis enzymes involved in the regulation of HS 6-*O*-sulfation in gastric cancer cells. (A) Illustrative image depicting the biosynthesis steps that modulate HS 6-*O*-sulfation. HS6STs catalyze the addition of sulfate groups in the C6 position of GlcNAc/GlcNS residues, during HS biosynthesis. HSULFs remove these 6-*O*-sulfate groups during post-synthesis editing reactions. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1). **(B)** The expression of *HS6STs 1-3* and *HSULFs 1-2* was determined for six gastric cancer cell lines via qRT-PCR. Bar graph shows the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. Two independent biological assays with technical triplicates were analyzed.

The MKN74 *HS6ST1* KO cell model was established *via* CRISPR-Cas9 gene editing, and biochemical characterization was performed *via* reversed-phase ion pair-high performance liquid chromatography (RPIP-HPLC)-based HS disaccharide analysis (Fig. 2A-C). We determined that all three 6-*O*-sulfated HS disaccharides present in the WT cells, monosulfated ΔHexUA-GlcNAc,6S (6S), disulfated ΔHexUA-GlcNS6S (NS6S) and trisulfated ΔHexUA,2S-GlcNS6S (NS6S2S) disaccharides, were not detected in the *HS6ST1* KO cells. In addition, we observed a trend suggesting increased levels of ΔHexUA-GlcNS (NS) and ΔHexUA,2S-GlcNS (NS2S) disaccharides, which might result from the accumulation of disaccharides that, in the WT condition, would be targeted by the 6-*O*-sulfotransferase (Fig. 2A and B). Noteworthy, structural analysis also showed a trend towards increased total levels of HS disaccharides in the *HS6ST1* KO cells (Fig. 2C).



HS sulfation (%)	WT	*HS6ST1* KO #2	*HS6ST1* KO #4	*HS6ST1* KO #5
NAc	50,79 ± 3,59	45,43 ± 2,10	47,90 ± 2,29	49,20 ± 2,81
NS	49,21 ± 3,59	54,57 ± 2,10	52,10 ± 2,29	50,80 ± 2,81
6OS	**16,71 ± 4,69**	**0,41 ± 0,70**	**0,00 ± 0,00**	**0,10 ± 1,17**
2OS	27,38 ± 5,09	28,48 ± 2,38	28,36 ± 3,54	27,49 ± 4,66
Total *OS*	44,09 ± 8,81	28,88 ± 1,74	28,36 ± 3,54	27,59 ± 4,51
Total *S*	93,30 ± 12,23	83,45 ± 3,89	80,46 ± 4,58	78,38 ± 7,24

Figure 2. *HS6ST1* KO cells display altered HS sulfation motifs. HS disaccharide composition and transcriptomic analysis of HS *O*-sulfation enzymes in *HS6ST1* KO gastric cancer cells was performed to characterize this cell model. **(A-C)** HS chains isolated and purified from MKN74 cells were digested with a mixture of Heparinases I, II, and III and analyzed by RPIP-HPLC to determine the structural composition and total amounts of HS disaccharides. **(A and B)** Relative quantities of distinct HS disaccharide units were measured. Bar graph shows the mean levels of each type of HS disaccharide: NAc = ΔHexUA-GlcNAc; NS = ΔHexUA-GlcNS; 6S = ΔHexUA-GlcNAc,6S; 2S = ΔHexUA,2S-GlcNAc; NS6S = ΔHexUA-GlcNS6S; NS2S = ΔHexUA,2S-GlcNS; 2S6S = ΔHexUA,2S-GlcNAc,6S; NS2S6S = ΔHexUA,2S-GlcNS6S. **(B)** Table represents the mean values of sulfation content. NAc = *N*-Acetylation; NS = *N*-Sulfation; 6S = 6-*O*-Sulfation; 2S = 2-*O*-Sulfation; total OS = total *O*-Sulfation, total S = total Sulfation (*N*- and *O*-sulfation). **(C)** Bar graph shows the mean levels of HS disaccharides, expressed in nanograms. Technical triplicates of biological duplicates were analyzed. **(D)** The expression of *HS2ST*, *HS6STs 2-3* and *HSULFs 1-2* was determined for six gastric cancer cell lines *via* qRT-PCR. Bar graph shows the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. Two independent biological assays with technical triplicates were analyzed.

Furthermore, to investigate if the KO of *HS6ST1* affected the endogenous expression of other enzymes involved in HS sulfation, the transcript levels of the sulfotransferases HS2ST1 and HS6ST 2-3, and sulfatases HSULF 1-2 were quantified through transcriptomic

analysis (Fig. 2D). qRT-PCR data showed that no significant consistent differences were observed when comparing the three *HS6ST1* KO clones with the parental cell line (Fig. 2D). Overall, these results reveal that HS6ST1 is the main isoform catalyzing HS 6-*O*-sulfation in MKN74 gastric cancer cells. Further, it shows that this modification can impact HS overall sulfation and disaccharide levels, suggesting the involvement of HS6ST1 in the regulation of HS biosynthesis.

As a complementary strategy to modulate HS 6-*O*-sulfation, we resorted to lentivirus transduction to establish MKN74 *HSULF2* overexpressing (OE) cell models. Recently, it was reported that HSULF2 is a CS/DS PG, whose enzymatic activity is modulated by the presence of this GAG chain. El Masri *et al.* showed that a mutant form of this enzyme lacking the GAG attachment site, HSULF2△SG OE, was associated with higher endosulfatase activity and pro-oncogenic property (323). It is currently hypothesized that HSULF2 can exist in both forms, with and without the GAG chain, and the ratio between these two existing forms may ultimately determine the outcome of this enzyme activity. Taking this information into consideration, we also established a cell model overexpressing the mutant endosulfatase without the CS/DS chain, MKN74 *HSULF2*△SG OE, to address the impact of both forms in HS 6-*O*-sulfation within the gastric cancer context.

Transcriptomic analysis confirmed the OE of *HSULF2* and *HSULF2*△SG and revealed that the expression of *HSULF1* remained unaffected in both cell models (Fig. 3A). The levels of HSULF2 in the transduced cells were also confirmed through Western blot (WB) analysis, using the 2B4 monoclonal antibody (mAb) that recognizes the enzyme C-terminal subunit (Fig. 3B). Although we could detect the *HSULF2* mRNA in MKN74 WT cells (Fig. 3A), its basal protein levels were too low to be detected *via* WB (Fig. 3B). In the transduced cell models, two sets of well-defined bands were detected: an intense band close to 100 kDa, corresponding the unprocessed full-length HSULF2, and a pair of bands between 37-50 kDa, corresponding to the C-terminal sub-unit, intact or partially degraded (Fig. 3B). Protein lysates were digested with Chondroitinase ABC (Chond. ABC), which catalyzes the digestion of CS chains, to address the GAGosylation profile of HSULF2 in MKN74 cells. No significant changes were spotted in the *HSULF2* OE protein cell lysates, suggesting that cellular HSULF2 (either being produced or attached to the cell surface after secretion) is mostly present in its GAG-free form. We also evaluated HSULF2 protein levels in another gastric cancer cell line, MKN45, expressing higher transcript levels of *HSULF2* (Fig. 1B), and confirmed the detection of HSULF2 by WB (Fig. 3B). Interestingly in this cell model, the labeling bands corresponding to the cleaved C-terminal fragments at 37-50 kDa were only detected upon the digestion of the protein lysates with Chond. ABC. HSULF2 cell-dependent PTMs (including GAGosylation), and consequent varying labeling profiles, have

been previously reported (323) and could justify the profile changes observed between different cell lines.

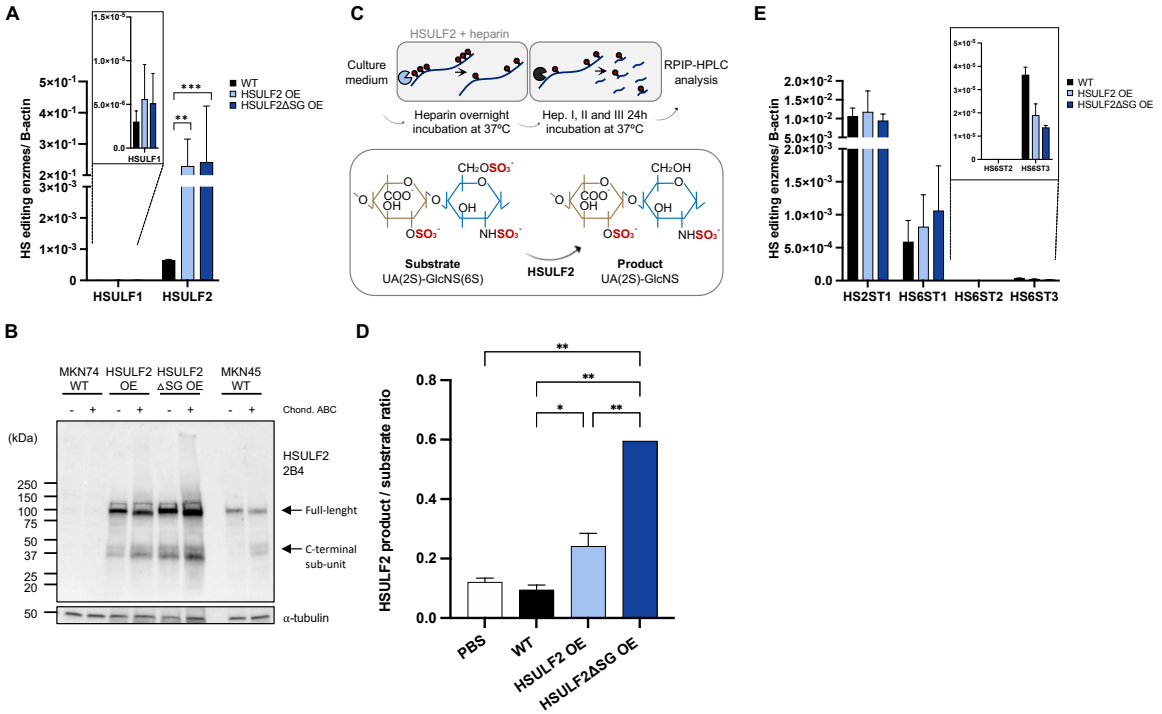


Figure 3. *HSULF2* and *HSULF2ΔSG* OE cells exhibit increased 6-*O*-endosulfatase activity. (A) The expression of *HSULFs 1-2* and **(E)** *HS2ST* and *HS6STs 1-3* was determined for *HSULF2* and *HSULF2ΔSG* OE cell models *via* qRT-PCR. Bar graphs show the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. Two independent biological assays with technical triplicates were analyzed. **(B)** MKN74 OE cell models and MKN45 WT cells were evaluated for HSULF2 total content by WB analysis. Protein lysates were immunoblotted for HSULF2 using the 2B4 mAb. Image is representative of two independent experiments. **(C and D)** HSULF2 enzymatic activity was evaluated *via* a 6-*O*-endosulfatase activity assay. Cell culture medium was incubated with heparin, following Heparinase I, II and III digestion and RPIP-HPLC analysis. HSULF2 main substrate and reaction product disaccharides were quantified. **(D)** Results are represented as the mean values of the product/substrate ratio + SD. Two independent experiments were analyzed (except for MKN74 *HSULF2*ΔSG OE cells, which were analyzed once).

To further characterize our cell models and verify the HSULF2 6-*O*-endosulfatase activity in the transduced cells, we performed an enzymatic assay by incubating the cell's conditioned medium (CM) with heparin, here used as a highly sulfated HS mimetic. Next, we performed RPIP-HPLC disaccharide analysis on untreated or CM-treated heparin and quantified the levels of HSULF2 main substrate and reaction product disaccharides: the trisulfated disaccharides UA(2S)-GlcNS(6S) and disulfated disaccharides UA(2S)-GlcNS, respectively (Fig. 3C and D). We determined that heparin treated with the CM from MKN74 *HSULF2* and *HSULF2*ΔSG OE cells showed significantly increased product/substrate ratio, indicating higher HSULF2 endosulfatase activity (Fig. 3C and D). Finally, transcriptomic analysis confirmed that the expression of HS 2-*O*-sulfation and 6-*O*-sulfation modification enzymes did not show any consistent alteration in these OE cell models (Fig. 3E).

HS6ST1 and HSULF2 modulate HS cellular content

Deligny *et al.* have previously established an association between the expression of *NDST2*, an enzyme that modulates HS *N*-sulfation motifs, and the degree of polymerization of HS chains in HEK 293 cells (391). Taking in consideration this possible HS sulfation-dependent regulatory mechanism that influences the length of HS growing chains, and aiming to assess the impact of the enzymes that modulate HS 6-*O*-sulfation in this mechanism, we further characterized the *HS6ST1* KO and *HSULF2* OE cell models, by evaluating the overall and cell surface levels of HS and HS stubs *via* WB and flow cytometry analyses.

For WB, protein lysates were incubated with Heparinase III (Hep. III), an enzyme that catalyzes the digestion of HS chains and generates smaller HS stubs capped with unsaturated GlcA residues. Non-digested and Hep. III digested protein lysates were immunoblotted for HS using the 10E4 mAb (Fig. 4A and B) and for HS stubs using the 3G10 mAb (Fig. 4C and D). Interestingly and in agreement with the trend observed in the HS structural analysis performed in these cell models (Fig. 2C), WB analysis revealed increased levels of HS in *HS6ST1* KO cells, as shown by the intense HS labeling smear (Fig. 4A). This variation in HS levels was not observed in *HSULF2* and *HSULF2*ΔSG OE cells (Fig. 4B).

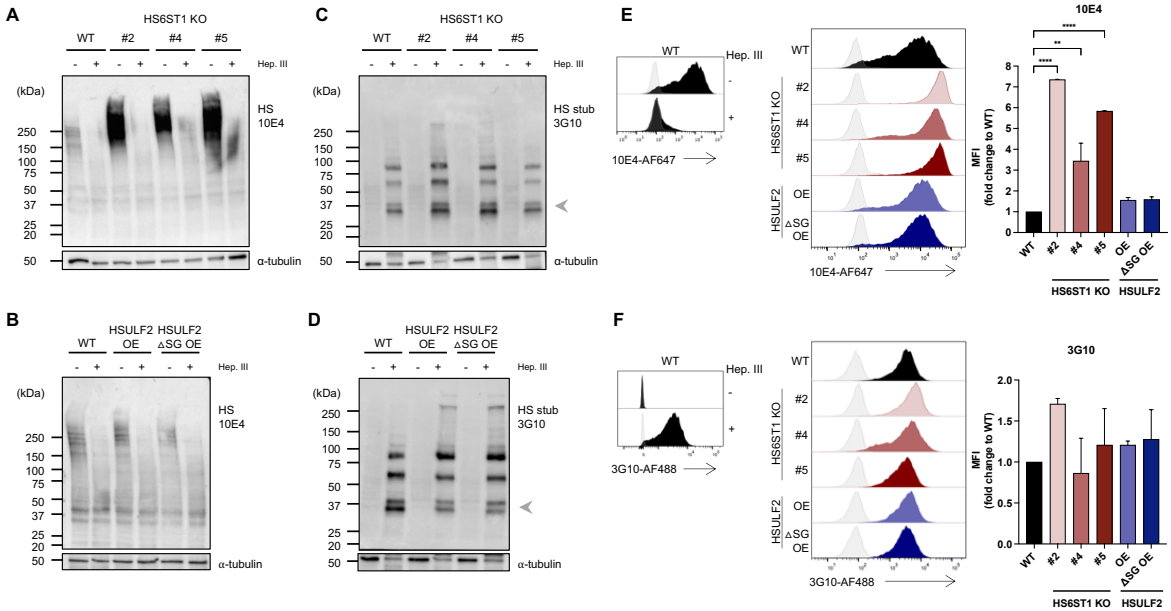


Figure 4. *HS6ST1* KO and *HSULF* 2 OE cells exhibit changes in total and cell surface levels of HS. (A-D) MKN74 *HS6ST1* KO, *HSULF2* and *HSULF2*ΔSG OE cells were evaluated for HS total content and number of HS chains by WB analysis. Protein lysates were treated with or without Heparinase III (Hep. III) and immunoblotted for **(A and B)** undigested HS using the 10E4 mAb or **(C and D)** HS stubs using the 3G10 mAb. Images are representative of two independent experiments. **(E and F)** The levels of cell surface HS and HS stubs were also quantified by flow cytometry analysis. Gray shaded peaks indicate the negative controls. The peaks colored in solid black, shades of red, or blue indicate the labeling with **(E)** 10E4 mAb or **(F)** 3G10 mAb of WT, *HS6ST1* KO, or *HSULF2* and *HSULF2*ΔSG OE cells, respectively. Bar graphs show normalized mean intensity of fluorescence (MIF) + SD. Images are representative of two independent experiments.

Regarding the labeling of HS stubs, as expected, several 3G10 signal bands were detected in the Hep. III digested protein lysates, corresponding to the cleaved HS chains modifying different HSPGs (Fig. 4C and D). Therefore, after the efficient Hep. III digestion, each signaling band should correspond to the labeling of a different HS-naked HSPG. When comparing the WT and *HS6ST1* KO cells, we observed two bands close to 37 kDa that were more intense in the KO condition (Fig. 4C), while the opposite trend was detected in both *HSULF2* and *HSULF2*△SG OE cells, which presented reduced 3G10 staining at the same molecular weight (MW) (Fig. 4D).

In accordance with the WB data, the flow cytometry analysis revealed significantly increased levels of HS in *HS6ST1* KO cells, and no significant variations in the *HSULF2* and *HSULF2*△SG OE cell models (Fig. 4E). In contrast, the 3G10 labeling did not vary significantly in all cell models (Fig. 4F). Unlike WB, flow cytometry analysis using the 3G10 mAb does not allow to distinguish opposite trends in the levels of different HSPGs within the same cell model, as this methodology results in the general quantification of 3G10 labeling. This means that increases in the level of a specific HSPG might cover up a decrease in the levels of other HSPG, which might explain the differences in HS stubs detection from WB and flow cytometry analyses. Additionally, it is possible that HS stubs from larger or heavily glycosylated HSPGs are not being detected.

HS6ST1 and HSULF2 impact cell glycoproteome

To understand the origin of the changes observed in HS and HS stub levels in *HS6ST1* KO and *HSULF2* OE cells, we performed qRT-PCR analysis and evaluated the expression of the HS biosynthesis enzymes involved in the initiation and polymerization stages, namely the glycosyltransferases Exostosin-Like 2 (EXTL2), EXTL3, Exostosin 1 (EXT1), and EXT2 (Fig. 5A). No significant changes were determined in the expression of these enzymes in any of the evaluated cell models. Additionally, we evaluated the expression of the main cell surface HSPGs expressed in the MKN74 cells (388): Syndecan-1 (SDC1), SDC4, Glypican-1 (GPC1) and GPC4 (Fig. 5B). Noteworthy, the expression of SDC4 was significantly higher in two of the *HS6ST1* KO cell clones. The overexpression of SDC4 in *HS6ST1* KO clones was confirmed *via* WB analysis, using the 8G3 mAb (Fig. 5C). These results not only evidenced a more intense SDC4 labeling smear in *HS6ST1* KO cells, when compared to the WT, but also revealed labeling of SDC4 at higher MWs, above 250 kDa (Fig. 5C). This suggests that in these glycoengineered cells, there is a subset of SDC4 modified with longer HS chains, and therefore larger. Interestingly, *HSULF2* and *HSULF2*△SG OE cells also presented increased labeling of SDC4 (Fig. 5C), which was not detect at the transcript level (Fig. 5B). Changes in SDC4 secretion and/or retention time at the cell surface might account for these differences. Protein levels of SDC1, the second highly expressed cell surface

HSPG in MKN74, were also evaluated *via* WB. Protein lysates were previously digested with Hep. III to allow detection with the anti-SDC1 B-A38 mAb (Fig. 5D). Although SDC1 transcription was unaltered in all cell models (Fig 5B), SDC1 protein levels were increased in *HS6ST1* KO and *HSULF2* and *HSULF2*ΔSG OE cell models, when compared to the WT (Fig. 5D). This observation further suggests that changing HS 6-*O*-sulfation might impact HSPGs cell surface retention or secretion.

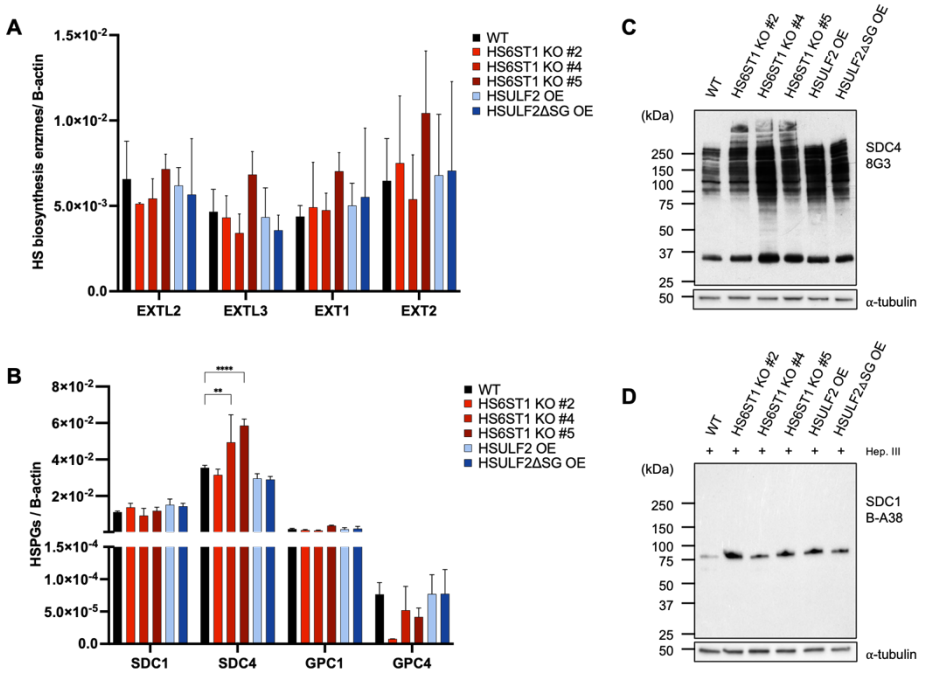


Figure 5. Modulation of HSPGs' expression and GAGosylation profiles by *HS6ST1* KO and *HSULF2* OE cells. The expression of **(A)** HS glycosyltransferases - Exostosin-Like 2 (*EXTL2*), *EXTL3*, Exostosin 1 (*EXT1*), and *EXT2* - and **(B)** cell surface HSPGs - Syndecan-1 (*SDC1*), *SDC4*, Glypican-1 (*GPC1*) and *GPC4* - was determined for the MKN74 *HS6ST1* KO, *HSULF2* and *HSULF2*ΔSG OE cells *via* qRT-PCR. Bar graphs show the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. Two independent biological assays with technical triplicates were analyzed. **(C)** SDC4 and **(D)** SDC1 protein levels were determined through WB. Protein lysates were treated without or with Heparinase III (Hep. III) for immunodetection of **(C)** SDC4 using the 8G3 mAb and **(D)** SDC1 using the B-A38 mAb.

HSULF2 is upregulated in gastric tumors and correlates with lower disease-free survival

To learn the expression profile and influence of HS6ST1 and HSULF2 in gastric tumors, we analyzed the transcriptomic profile of both enzymes using samples from a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD) (Fig. 6A-F). We learned that while *HS6ST1* expression did not vary (Fig. 6A), *HSULF2* expression was significantly increased in gastric tumors from the intestinal and diffuse subtypes (Fig. 6D). Kaplan–Meier (KM) analysis of *HS6ST1 and HSULF2* mRNA levels and gastric cancer patients' overall survival (OS) and disease-free (DFS) was performed using the same TCGA-STAD cohort (Fig. 6B-C and E-F). This analysis further revealed a

significant correlation between *HSULF2* expression and worse prognosis (lower DFS; p-value=0,013) for the patients with gastric cancer from the intestinal subtype (Fig. 6E). Moreover, although not statistically significant, we also observed a trend associating *HS6ST1* expression with lower OS (p-value=0,087) of gastric cancer patients included in the intestinal subtype category (Fig. 6C).

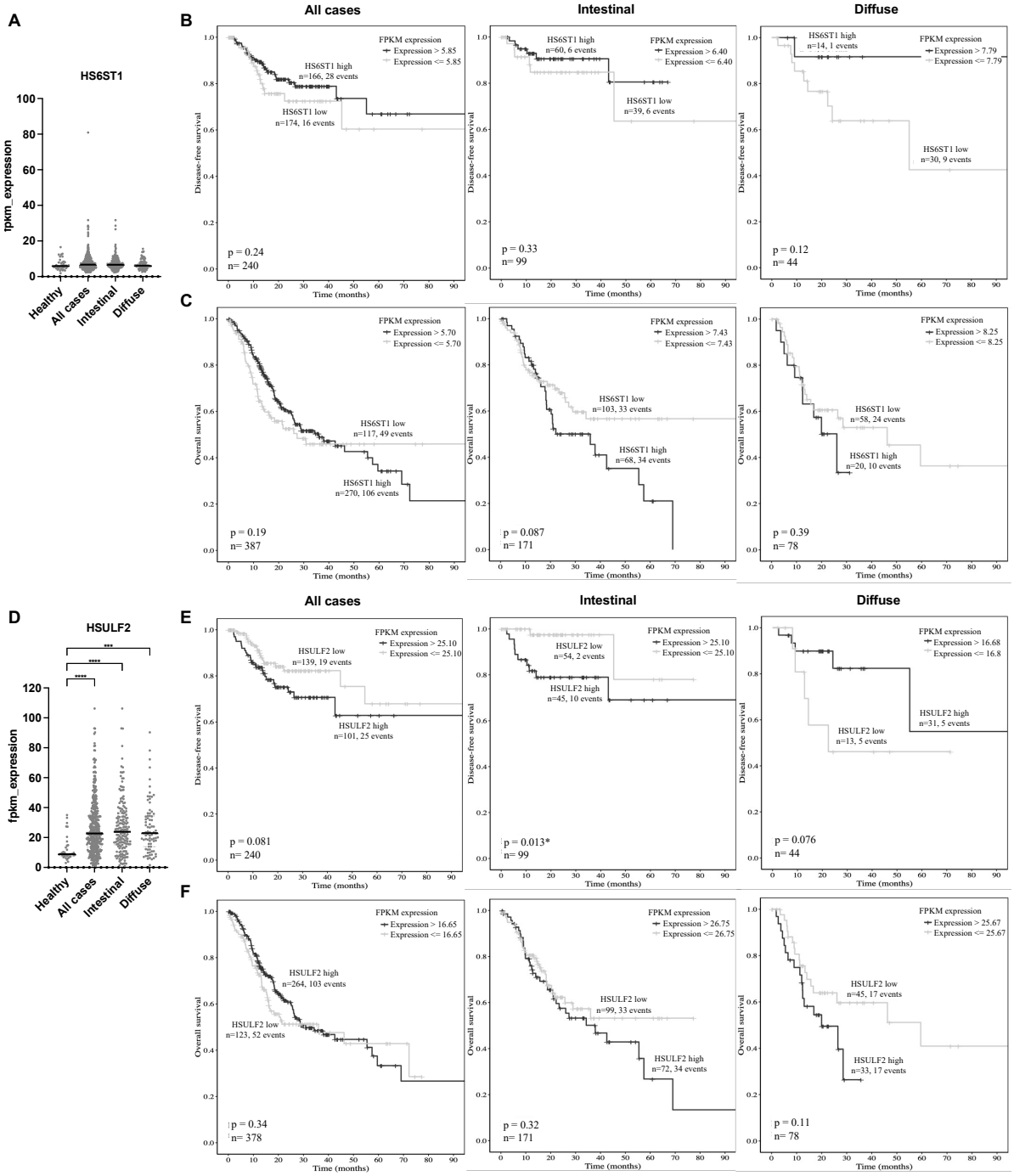


Figure 6. *HSULF2* is highly expressed in intestinal subtype gastric cancer and it associates with patients' lower disease survival. Distribution of **(A)** *HS6ST1* and **(D)** *HSULF2* mRNA levels in a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD), according to the Lauren classification. mRNA gene expression data was quantified and represented as FPKM (fragments per kilobase million). Disease-free survival (DFS) and overall survival (OS) associated to **(B and C)** *HS6ST1* and **(E and F)** *HSULF2* gene expression were assessed *via* Kaplan-Meier (KM) analysis using the same TCGA-STAD patient cohort.

DISCUSSION

HS 6-*O*-sulfation patterns are regulated by HS6STs and HSULFs at different stages of the HS biosynthetic and post-synthetic pathways, foreshadowing the biological relevance of this specific structural feature in homeostasis. Several studies have detailed how these enzymes are abnormally expressed in the context of cancer, further highlighting the dynamic nature of HS 6-*O*-sulfation epitopes as a potential player in carcinogenesis. Understanding how these enzymes modulate HS and HSPGs biosynthesis and cellular distribution, and ultimately, the impact of these changes in tumor cell behavior is fundamental and it might set a steady foundation for the development of new diagnostic and therapeutic tools.

The sulfotransferases HS6STs are described as HS enzymes with broad substrate recognition and enzymatic activity, and despite the conflicting data on the preferred substrate for each isoform, there is a general consensus on the ability of all isoforms to transfer sulfate groups to GlcNAc and GlcNS residues, preferably on IdoA-containing disaccharides (164-166). In our study, HS disaccharide analysis showed that although both *HS6ST1* and *HS6ST3* were expressed at detectable levels on MKN74 cells (Fig. 1B), HS6ST1 was the primary isoform catalyzing HS 6-*O*-sulfation, as *HS6ST1* KO abolished the synthesis of all HS 6-*O*-sulfated disaccharides (Fig. 2A). Additionally, *HS6ST1* KO indirectly influenced the levels of other *N*- and *O*-sulfated motifs. Structural analysis revealed increased levels of NS and NS2S disaccharides, which might have resulted from the accumulation of disaccharides that were not further modified into NS6S and NS2S6S, respectively, due to the absence of *HS6ST1*. Other reports have previously described that 6-*O*-sulfation of GlcNAc/GlcNS residues might preclude the activity of HS2ST1 towards the adjacent IdoA residues (162, 168). This might also explain the increased levels of NS2S disaccharides in the absence of *HS6ST1* expression. Nonetheless, we did not detect changes in the levels of monosulfated 2S disaccharides (Fig. 2A and B).

HSULFs are secreted extracellular endosulfatases with catalytic activity towards cell surface and extracellular HSPGs. These enzymes have a more specific substrate recognition, targeting primarily the trisulfated NS2S6S disaccharides at the non-reducing end of the highly sulfated S-domains (185, 186). In our work, we could not detect HSULF2 protein levels (Fig. 3B) or extracellular enzymatic activity (Fig. 3D) in MKN74 WT cells, despite the detectable mRNA transcript levels (Fig. 3A). Upon OE of *HSULF2* and *HSULF2*ΔSG, we observed increased endosulfatase activity (Fig. 3D). Similarly to what has been previously reported by El Masri *et al.*, this increase was higher in cells OE *HSULF2* lacking the CS/DS chain (323). Digestion of protein lysates with Chond. ABC for WB analysis confirmed, as expected, the absence of CS/DS chain for HSULF2ΔSG OE and revealed that HSULF2 is mostly present in its GAG-free form, when overexpressed in

MKN74 cells (Fig. 3B). Noteworthy, this analysis only detects cell-associated HSULF2. As HSULF2 GAG chain reduces binding of the enzyme to the cell surface, it would be interesting to also analyze HSULF2 present in the secretome of these cells, to accurately evaluate the ratio between GAG-free and GAG-modified enzyme and better understand its functional role in the context of gastric cancer. In addition, evaluation of HSULF2 levels in MKN45 cells, here used as a positive control for the 2B4 labeling, suggested a higher prevalence of the GAG-modified form in these cells (Fig. 3B).

In previous work, it was reported that HS sulfation influences GAG chain polymerization. In a former study, *N*-sulfation of a terminal GlcNAc in a nascent HS chain was shown to stimulate the incorporation of GlcA residues, further stimulating chain elongation (392). More recently, NDST2 from the family of NDST enzymes, which are mainly responsible for GlcNAc *N*-deacetylation and sulfation, playing an important part in the regulation of HS overall structure and sulfation, was specifically highlighted for its ability to promote HS polymerization in HEK 293 cells (391). Our results revealed that abrogation of HS6ST1 induced HS synthesis (Fig. 2C and Fig. 4), although the expression of HS glycosyltransferases remained unaltered (Fig. 5A). This suggests that the rate of HS polymerization is influenced by HS6ST1 activity, rather than by compensatory mechanisms involving the expression of HS-associated enzymes. We theorize that, similarly to what has been proposed for NDST2, HS6ST1-mediated changes in the HS growing chains, either null 6-*O*-sulfation or increased *N*-sulfation or *N*- and 2-*O*-sulfation, might be targeted by the polymerizing enzymes. On the other hand, transcriptomic data also showed significantly increased expression of the cell surface HSPG SDC4, which could also account for the increased HS levels (Fig. 5B). Nonetheless, WB analysis revealed the synthesis of higher MW SDC4 molecules in the absence of HS6ST1 (Fig. 5C), suggesting the presence of SDC4 modified with longer HS chains, which would be aligned with the forementioned data. It would be important to confirm these data and address specifically the polymerization of HS chains in the *HS6ST1* KO cell models, which could be achieved through metabolic labeling and gel chromatography. It would also be important to evaluate the labeling profile of other HSPGs besides SDC4 to understand if this is a general mechanism affecting cells glycoproteome or if the HS6ST1-mediated effects on HS polymerization are specifically targeting SDC4 in the MKN74 cells. Although we also evaluated SDC1 protein levels (Fig. 5D), the need of digesting HS for anti-SDC1 mAb efficient labeling was a limitation to infer on the HSPG GAGosylation. Finally, it was interesting to note that although both *HS6ST1* KO and *HSULF2* OE had similar effects on HS 6-*O*-sulfation (possibly at a different extent), the significant effects on HS levels were only detected in the absence of HS6ST1, which further confirms that these changes might occur intracellularly during HS biosynthesis, rather than at the extracellular level through post-synthesis mechanisms.

Lai *et. al.* previously outlined an association between *HSULF2* and *GPC3* expression with patients' worse prognosis, in the context of hepatocellular carcinoma. They reported that overexpression of *HSULF2* in the human liver cancer cell line Hep3B induced the upregulation of *GPC3*, a relevant HSPG in this pathology, with possible consequences in cell proliferation and migration (322). In our study, when addressing the impact of altered 6-*O*-sulfation in HSPG expression and cellular protein levels, we observed that abrogation of *HS6ST1* promoted SDC4 expression and changed its GAGosylation profile. On the other hand, *HSULF2* OE had no impact in any of the SDCs or GPCs transcripts evaluated but led to a slight increase in SDC4 protein levels (Fig. 5C). Noteworthy, we have recently reported that *SDC4* is upregulated in gastric cancer (associated with the intestinal subtype), associated with patients' poor survival, and revealed its significant role cell motility and invasion (270). Therefore, our results suggest a possible new mechanism for the regulation of *SDC4* expression with impact in the gastric cancer context. Additionally, altered expression of either *HS6ST1* or *HSULF2* appeared to induce equally the increase of cellular levels of SDC1, with no impact over its expression, suggesting that decreased HS 6-*O*-sulfation might contribute for increased half-life of SDC1 in MKN74 cells (Fig. 5D). New experiments would have to be performed to address this theory, including flow cytometry analysis to quantify cell surface levels of HSPG and evaluation of cells' secretome to evaluate possible changes in HSPG shedding. Overall, these results highlight the imperative need to evaluate the impact of these changes in HS and cell glycoproteome in cell signaling, testing the activation of classical ligand-receptor axis known to be influenced by HS 6-O-sulfation motifs.

Data on the clinical impact and role of HS 6-*O*-sulfotransferases and endosulfatases in gastric cancer is quite scarce and conflicting. *HS6ST2* upregulation has been associated with patients worse prognosis (331), *HSULF1* has been associated with both pro- (335) and anti-oncogenic roles (334), while *HSULF2* overexpression has been reported in gastric malignancies (335). Here, we report that *HSULF2* is overexpressed in gastric cancer from the intestinal and diffuse subtypes, and that it associates with patient's poor disease-free survival (Fig. 6). Functional analyses are required to further understand how the specific modulation of HS 6-*O*-sulfation by HSULF2 promotes this more aggressive phenotype and to uncover the associated signaling cascades.

In conclusion, in this study we disclosed that the sulfotransferase HS6ST1 is the main isoform regulating HS 6-*O*-sulfation in gastric cancer MKN74 cells during biosynthesis, and revealed its general influence over HS sulfation motifs. Noteworthy, we showed that intracellular (*HS6ST1* KO) or extracellular (*HSULF2* OE) changes in HS sulfation influence differently HS cellular levels and expression and cellular presentation of main cell surface HSPGs. Finally, our results showed that the endosulfatase *HSULF2* is highly expressed in

gastric cancer and it associates with patients' worse prognosis, highlighting the clinical potential of this enzyme in gastric cancer therapy.

EXPERIMENTAL PROCEDURES

Cell lines, cell culture and genetic engineering

In this study, we resorted to the human gastric adenocarcinoma cell line MKN74, obtained from the Japanese Collection of Research Bioresources Cell Bank. Cells were cultured in RPMI–1640 (Gibco) culture medium, supplemented with 10% (v/v) Fetal Bovine Serum (FBS) (Biowest), at 37°C, under 5% (v/v) CO_2 conditions.

The glycoengineered *HS6ST1* KO cell model was generated from the gastric cancer cell line MKN74, *via* CRISPR-Cas 9 genome editing. Briefly, cells were co-transfected with Cas9 endonuclease vector containing GFP and a plasmid with a validated gRNA for the target gene (Table 1). KO clones were obtained by single cell sorting of FACS enriched nuclease-expressing cells and gene KO was validated by indel detection of PCR products by fragment analysis, as previously described (383). From fragment analysis three clones presenting distinct indels were selected for each target gene and mutations confirmed by Sanger sequencing. Analysis was performed through Tracking of Indels by Decomposition (TIDE) methodology (384).

Table 1. *List of validated gRNAs used to target HS6ST1 and results of indel analysis.*

Gene	gRNA sequence	Indels		
HS6ST1	GGAGCTGCCGCCCTGCTACG	clone 2	clone 4	clone 5
		+1	-11; +1	-20

To establish *HSULF2* OE cell models (WT and ΔSG mutant), lentiviruses production and lentiviral transduction was performed as previously described in (323). Briefly, cells were plated into 6 well-plates ($4x10^5$ in 3 mL of serum-supplemented medium). A day later, cells were incubated with 1 mL of lentiviral medium plus 1 mL of 16 µg/mL of polybrene (Sigma/Aldrich) diluted in cell medium for 4h, after which an additional mL of supplemented medium was added. Cells were maintained approximately for 24h before washing. For stable expression, the cells were then maintained in puromycin supplemented medium (2 µg/mL, Life Technologies).

Antibodies

The antibodies and conditions used for the described analyses are listed in Table 2.

***Table 2**. List of antibodies and conditions used for WB and Flow Cytometry (FACS) assays.*

Antibodies	Applications and Dilutions		Source/Reference
	WB	FACS	
HS (F58-10E4)	1:600	1:600	Amsbio
HS stub region (F69-3G10)	1:500	1:200	Amsbio
SDC1 (B-A38)	1:125		Abcam
SDC4 (8G3)	1:2000		(386)
HSULF2 (2B4)			R&D Systems
α-Tubulin (DM1A)	1:10000		Sigma-Aldrich
Horseradish Peroxidase-conjugated α mouse IgG	1:5000		Jackson ImmunoResearch Laboratories, Inc
Horseradish Peroxidase-conjugated α mouse IgG1	1:8000		Jackson ImmunoResearch Laboratories, Inc
Horseradish Peroxidase-conjugated α mouse IgM	1:5000		Jackson ImmunoResearch Laboratories, Inc
Alexa Fluor® 647 α mouse IgM		1:400	Jackson ImmunoResearch Laboratories, Inc
Alexa Fluor® 488 α mouse IgG		1:200	Invitrogen

HS detection by Flow Cytometry Assay

Cells previously cultured in 75cm^2 flasks, were detached with accutase (GRISP), resuspended in fresh medium and centrifuged at 300g for 5 min. For HS stub analysis, cells were previously digested with Heparinase III (EC 4.2.2.8; 5 mU/mL in TBS 0.1% Bovine Serum Albumin (BSA); Amsbio) in RPMI supplemented with 2% FBS and 2mM $CaCl_2$, for 1h at room temperature. Cells (1×10^6 cells/mL) were washed 2 times in Phosphate Buffered Saline (PBS), 1% BSA (Sigma-Aldrich) solution, each wash followed by centrifugation at 300g for 5 min. The cell pellet was resuspended and incubated 1h at room temperature with the primary antibody 10E4 or 3G10 (Table 2) in PBS, and then washed again with PBS 1% BSA solution and incubated with Alexa Fluor® 647 α mouse IgM or Alexa Fluor® 488 α mouse IgG. For negative controls, cell pellets were incubated with secondary antibodies only. Lastly, cells were washed again and resuspended in PBS 1% BSA solution and filtered for analysis. Data was acquired using FACscanCanto II and analyzed with the FlowJo Software (v10). Mean Intensity of Fluorescence (MIF) values measured for the KO clones were normalized to the MIF values of WT, which were defined as a unit value. Two independent experiments were analyzed.

Western Blotting

Total protein lysates were obtained from confluent cells scraped and lysed with lysis buffer 17 (R&D Systems) supplemented with 1 mM sodium orthovanadate (Sigma-Aldrich), 1 mM phenylmethanesulfonyl fluoride (Sigma-Aldrich) and cOmplete™ protease inhibitor

cocktail (Roche). The protein concentration of these lysates was determined using the DC protein assay (BioRad). Total protein lysates (30 µg for HS, 12,5 µg for HS stub, 25 µg for SDC1, SCD4 and HSULF2) were incubated with 4x Laemmli Sample Buffer (BioRad), supplemented with 10% β-mercaptoethanol, at 98ºC for 5 min. The protein samples were loaded and run in 4-15% precast polyacrylamide gels (BioRad) and blotted to a nitrocellulose membrane (GE Healthcare Life Sciences). The membranes were blocked for 1h at room temperature, with TBS 5% BSA and 0.05% Tween® 20 (Sigma-Aldrich) (TBS-T 0.05%) buffer, and then incubated overnight at 4°C with primary antibodies (Table 2) diluted in the blocking solution. Membranes were washed with TBS-T 0.05% and incubated for 1h at room temperature with horseradish peroxidase-conjugated secondary antibodies diluted in the blocking solution. New washes with TBS-T 0.05% were performed, membranes were developed, and the protein bands visualized with ECL chemiluminescent western blotting detection reagent and films (GE Healthcare Life Sciences). Two independent experiments were analyzed.

Glycosaminoglycan Enzymatic Digestion Assay

HS enzymatic digestion was performed by incubating 250 µg of protein lysates of each sample with Heparinase III (EC 4.2.2.8; 5 mU/mL; Amsbio) in TBS 0.1% BSA, supplemented with Calcium Acetate (0.1 mM), overnight, at 37ºC with continuous mixing. 25µg of the resulting digested samples were then analyzed through WB analysis, by labeling HS digested resulting stubs with 3G10 antibody (Table 2). CS enzymatic digestion was performed by incubating 50 µg of protein lysates of each sample with Chondroitinase ABC (EC 4.2.2.4; 10 mU/µl; Amsbio) in TBS 0.1% BSA, supplemented with 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$, overnight, at 37ºC with continuous mixing.

Real-time PCR analysis

For mRNA quantification/gene expression analysis, total RNA was extracted from MKN74 WT and glycoengineered cell models (*HS6ST1* KO, *HSULF2 and HSULF2* ΔSG OE) using TRIzol Reagent (ABP Biosciences). 5 µg of RNA from each cell condition was converted into cDNA *via* reverse transcription by incubation with random hexamers/primers, using the SuperScript™ IV Reverse Transcriptase Kit (Invitrogen). Real-time PCR was performed using for each condition, per well: 4 µL of cDNA diluted 1:10 in ultrapure water, 0.6 µL of each primer at 10 µM (Table 3), 10 µL of PowerUp™ SYBR™ Green Master Mix (Thermo Fischer Scientific) and ultrapure water up to a final volume of 20 µL. The real-time PCR run was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems). RQ values were determined for each gene and 18S ribosomal RNA was used as a

housekeeping gene to normalize relative gene expression (RQ = $2^{-(\text{target Ct} - \text{18S Ct})}$). Two independent experiments with triplicates were analyzed.

Table 4. *List of primers used in the real-time PCR analysis.*

Gene	Primer Fw	Primer Rv
SDC1	5'-ATGGCTCTGGGGATGACTCT-3'	5'-GTGGGAATAGCCGTCAGGAG-3'
SDC4	5'-CCGGAGCCCTACCAGACGAT-3'	5'-AGGCACCAAGGGATGGACAA-3'
GPC1	5'-CATCGGGTGTGGAGAGTG-3'	5'-TGAGCGTGTCCCTGTTGTC-3'
GPC4	5'-AGTGTGGTCAGCGAACAGTG-3'	5'-CAAACATATCATTCAGGGATTTCTC-3'
EXTL2	5'-TGAACTGGAAACCAATGCAG-3'	5'-AGGAAATTGCTGCCAAACTG-3'
EXTL3	5'-CTCCGCCATCAGGAAATC-3'	5'-AGTTGGAGTTGTAGAGCCAGGA-3'
EXT1	5'-GCAGGCTTGGGTCCTTCAGAT-3'	5'-CATCGCCTATGACGGCAGCT-3'
EXT2	5'-AGGACCTAGAAGCCCTCCAG-3'	5'-GAGCTCCACGAAGAACCACA-3'
HS2ST1	5'-TGGAGATGATTATAGACCAGGGTTAC-3'	5'-GCTATGGCCACAGAAGAACG-3'
HS6ST1	5'-GCAGGGAGTGGAGCTAACAG-3'	5'-AACAGTTCCAGTTCCCGAAA-3'
HS6ST2	5'-CGGTGCGATCTTCTCCAA-3'	5'-AGGACGATCACGGCAAATAG-3'
HS6ST3	5'-CAACCACAGCCACACCAG-3'	5'-CTTCTTCCATCACACATATGAAGAG-3'
HSULF1	5'-CCAGCAGAAGCCAAAGAAAG-3'	5'-GAACGTGTCTGCCGAGTATG-3'
HSULF2	5'-GCCTGCAAGAGAAGGACAAG-3'	5'-AGCAGCTTGCGGAGTTTC-3'
B-actin	5′-TGCTATCCAGGCTGTGCTAT-3′	5′-AGTCCATCACGATGCCAGT-3′

GAG purification and HS/CS disaccharide analysis

To evaluate HS disaccharide composition, GAGs were isolated and purified, and structural analysis was performed as described in (393). For preparation and purification of GAG chains, cells were exhaustively digested with trypsin (45 min at 37°C). Cellular debris were discarded by centrifugation and supernatants were recovered and incubated with 250 U/µL Benzonase (Merck) and 2 mM $MgCl_2$ for 2 h at 37°C. Benzonase was then inactivated by heating supernatants at 96°C for 5 min and additional centrifugation was performed to discard nucleic acids. The recovered supernatants were applied to a DEAE-Sephacel column (Pharmacia Biotech) equilibrated in 20 mM phosphate pH 6.5. The column was then extensively washed with 20 mM phosphate pH 6.5, 0.3 M NaCl, then GAGs were eluted with 20 mM phosphate pH 6.5, 1 M NaCl. Samples were desalted over a Pd-10 column (GE Healthcare), lyophilized, and stored at -20°C prior to analysis. For HS disaccharide analysis, samples were resuspended in 100 mM sodium acetate, 0.5 mM calcium acetate, pH 7.1, then digested into disaccharides by incubation with a mix of heparinase I, II and III (10 mU each) for 48 h at 37°C. Compositional analysis were performed by RPIP-HPLC, by applying samples to a C18 reversed phase column equilibrated in a H_2O/acetonitrile (8.5%) buffer supplemented with 1.2 mM of ion-pairing tetra-*N*-butylammonium hydrogen sulfate (TBA), then resolved using a multi-step NaCl gradient calibrated with HS disaccharide standards. On-line post-column disaccharide derivatization was achieved by the addition of 2-cyanoacetamide (0.25%) in NaOH (0.5%), followed by fluorescence detection (excitation 346 nm, emission 410 nm).

HSULF2 enzymatic activity assay

Evaluation of HSULF2 enzymatic activity was performed following the protocol previously described in (394). Conditioned medium was collected from highly confluent cells cultured in 75cm^2 flasks and incubated overnight with heparin at 37ºC. This was followed by incubation with a mixture of Hep. I, II, and III (10 mU each). The processed samples were then subjected to RPIP-HPLC based HS disaccharide analysis, as described before. For analysis, we calculated the ration between the quantified levels of the disulfated (UA(2S)-GlcNS) and trisulfated (UA(2S)-GlcNS(6S)) disaccharides (product/substrate). Results were represented as the mean values of the product/substrate ratio + SD. Two independent experiments were analyzed (except for MKN74 HSULF2 ΔSG cells, which were analyzed once).

TCGA data analysis

mRNA gene expression data, quantified as FPKM (fragments per kilobase million) was retrieved from the Cancer Genome Atlas (TCGA) and comprised a total of 410 stomach adenocarcinoma samples and 38 normal gastric tissues. Clinical information from these samples was also downloaded from TCGA. Information on recurrence was downloaded from cBioPortal (https://www.cbioportal.org/).

Gene expression of *HS6ST1* and *HSULF2* was evaluated according to the available clinicopathological data, namely Lauren classification, following the information detailed in table 3. The "all cases" category included the data from *diffuse*, *intestinal* and *other* subtypes.

Table 3. *Gastric tumor tissue categories included in each subtype.*

Diffuse	Stomach, Adenocarcinoma, Diffuse Type
	Stomach Adenocarcinoma, Signet Ring Type
Intestinal	Stomach, Intestinal Adenocarcinoma, Not Otherwise Specified (NOS)
	Stomach, Intestinal Adenocarcinoma, Papillary Type
	Stomach, Intestinal Adenocarcinoma, Mucinous Type
	Stomach, Intestinal Adenocarcinoma, Tubular Type
Other	Stomach, Adenocarcinoma, Not Otherwise Specified (NOS)

Overall survival (OS) and disease-free survival (DFS) analysis were assessed using the KM test in R using the "survival" and "survminer" packages. Only individuals with survival information above zero were considered. Individuals were clustered into "high" and "low" expression groups based on the best separation cut-off method. This cut-off refers to the value, within the second quartile of expression of a given gene, that yields the lowest log-rank P-value. Only comparison with a log-rank P-value below 0.05 were considered.

Statistical Analysis

Statistical analysis was carried out using GraphPad Prism 9 and statistical significance was considered when P values were ≤0.05 (* means p≤0.05; ** means p≤0.01; *** means p≤0.001; **** means p≤0.0001). For flow cytometry, statistical significance of WT versus *HS6ST1* KO, *HSULF2* and HSULF2 ΔSG OE was determined using the ordinary one-way ANOVA, with a 95% interval of confidence. For GAG disaccharide structural analysis and real-time PCR, statistical significance was determined with the one-way ANOVA using Tukey's test for multiple comparisons. In regards to the TCGA analysis, the statistical significance of the mRNA gene expression data was determined using ordinary one-way ANOVA with Tukey's multiple comparisons test.

SUPPORTING INFORMATION

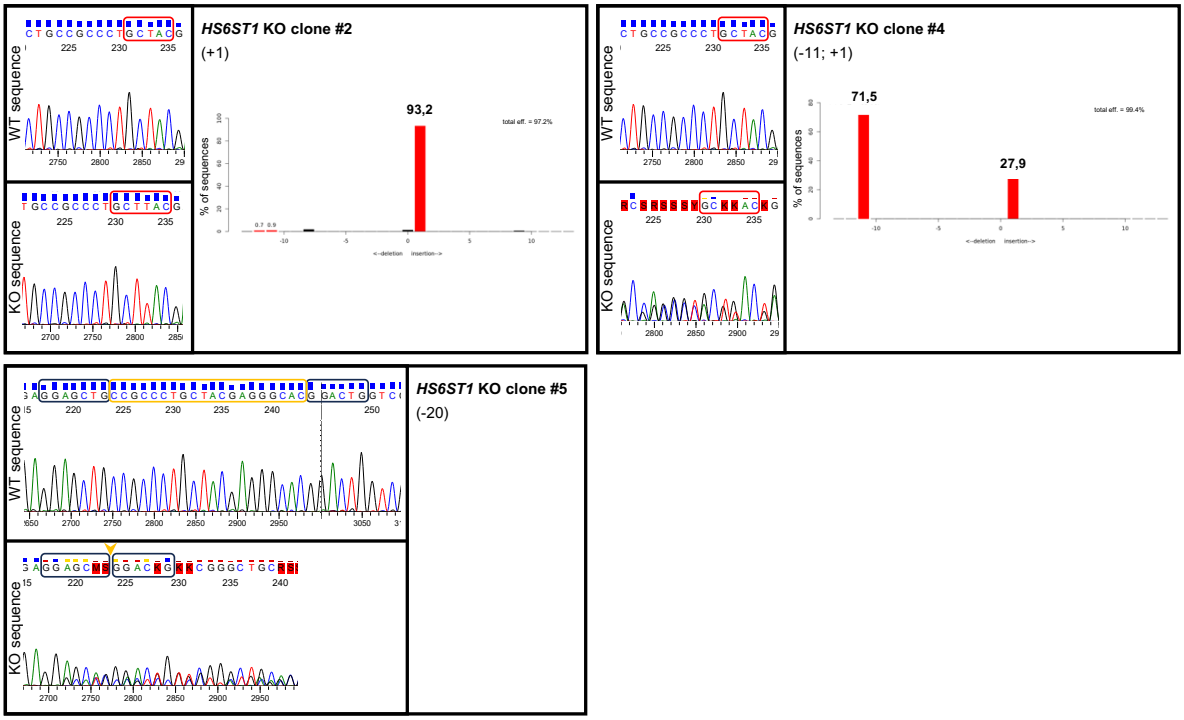


Figure S1. Validation of the glycoengineered *HS6ST1* KO cell models generated via CRISPR-Cas 9 genome editing. KO validation through Sanger sequencing for *HS6ST1* KO clones #2, #4 and #5 Indel sequencing was validated through Tracking of Indels by Decomposition (TIDE) methodology. The MKN74 WT sequence (top left panel) was used as the reference (control) for comparison with all KO clones. For *HS6ST1* KO clone #4 gene ablation resulted from a biallelic mutation (-11; +1), hence Sanger sequencing data showed two different readings for several nucleotides (consistent double peaks). The nucleotide insertion (+1) was highlighted in the Sanger sequencing graph. TIDE analysis validated the biallelic mutation and gene KO. For *HS6ST1* KO clone #20, the Sanger sequencing data suggested that gene KO resulted from a biallelic mutation (consistent double peaks suggestive of two reads). Nucleotide deletion (-20 indel) was assessed through fragmentation and sanger sequencing analysis (*yellow box* and *arrowhead* indicate the 20 deleted nucleotides and the region of deletion in clone #5, respectively; *black boxes* highlighted similar sequences). TIDE analysis was not performed as the limited coverage of decomposition did not allow the indel detection.

Chapter 4

Identification of a gastric cancer-associated GAGosylation signature

Uncovering the Gastric Cancer Glycosaminoglycan Landscape

Catarina Marques [1,2,3,4] Ana Raquel Ribeiro [1,2,6], Filipe Pinto[1,2], Bruno Cavadas[1], Rita Barros [1,2,7], Fátima Carneiro [1,2,6,7,] Celso Reis [1,2,5,6,] Romain Vivès [4,] Ana Magalhães [1,2,5*]

1 - i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto 4200-135, Portugal.

2 - IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto 4200-465, Portugal.

3 - Programa Doutoral em Biologia Molecular e Celular (MCbiology), Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal.

4 - Univ. Grenoble Alpes, CNRS, CEA, IBS, Grenoble, France.

5 - ICBAS - Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto 4050-313, Portugal.

6 - FMUP - Faculdade de Medicina da Universidade do Porto, Porto 4200-319, Portugal.

7 - Department of Pathology, Centro Hospitalar Universitário São João, 4200-319 Porto, Portugal.

[*] Corresponding author: amagalhaes@ipatimup.pt

Manuscript in preparation

ABSTRACT

Gastric cancer represents an important health concern, due to its aggressive profile and the limited availability of effective treatment and detection strategies. Identifying the molecular/phenotypic changes taking place during gastric carcinogenesis might provide additional insight on the molecular mechanisms that underly tumor formation and support the development of novel clinical approaches. Proteoglycans (PGs) and glycosaminoglycans (GAGs) are key molecules, endowed with important structural and biochemical functions within cells glycocalyx and the extracellular matrix (ECM). These molecules have been implicated in cancer development and progression, as well as in the modulation of the tumor microenvironment. Changes in the content of PGs and GAGs, along with alterations in GAG structural features, take place during the carcinogenic process, and associate with tumor cells malignant transformation, emphasizing the clinical potential of these molecules in cancer therapy.

Initially, we analyzed the transcriptomic profile of Heparan Sulfate (HS) and Chondroitin Sulfate (CS)/Dermatan Sulfate (DS) enzymatic machinery and main cell surface HSPGs, using six gastric cancer cell lines with different histopathological origin. qRT-PCR analysis revealed that all gastric cancer cell lines expressed the enzymatic machinery to carry all major steps of HS and CS/DS biosynthesis. Moreover, Syndecan-1 (SDC1), SDC4 and Glypican-1 (GPC1) were the main cell surface HSPGs detected in all tested cell lines. To identify possible tumor-associated HS and CS/DS transcriptomic profiles, transcriptomic analyses were employed using samples from a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD). These data outlined an interesting trend suggesting overall increased HS and CS formation, and decreased HS 6-*O*-sulfation in gastric malignancy, in both diffuse and intestinal subtypes. Additionally, to validate the clinical relevance of specific HS and CS/DS profiles, we optimized a methodology for GAG disaccharide structural analysis using gastric cancer patient tissue samples.

Overall, our work outlined the cell-type specificity and the importance of GAGs and GAG biosynthesis in gastric cancer, and further supported that deregulation of the HS/CS metabolic system impacts cancer progression and patient prognosis.

INTRODUCTION

Gastric cancer poses a concerning health challenge, currently ranking as the fifth most commonly diagnosed cancer and fifth leading cause of cancer related deaths worldwide (2). The asymptomatic nature of this disease during progression delays its diagnosis until later stages, which combined with the current limited treatment strategies, underlies the aggressiveness of this malignancy and contributes to the associated high mortality (24).

The need to develop new specific and sensitive diagnostic tools and efficient therapeutic options has become increasingly evident. Several studies exploring the role of glycosylation in tumor malignant transformation have emerged to fulfill this need (51, 52). Deregulated glycans biosynthetic pathways and aberrant glycans expression have been largely reported as main driving features of disease onset and progression, and shown to influence tumor cell signaling cascades, cell proliferation and death evasion, invasion and metastasis, angiogenesis and immune escape (52, 54-56). Proteoglycans (PGs) have been highlighted as glycoconjugates of interest in cancer pathology.

PGs entail an important and abundant class of glycoconjugates present in cells' glycocalyx and extracellular matrix (ECM). PGs participate in multiple signaling pathways associated to cell growth and proliferation, formation and distribution of extracellular gradients, cell adhesion, migration and invasion, membrane trafficking, and angiogenesis, contributing to many biophysical and biochemical events both in homeostasis and pathophysiology (84, 85).

PGs are composed by a core protein with at least one covalently attached glycosaminoglycan (GAG) chain. There are different classes of PGs that differ from each other depending on the core protein, cellular and subcellular localization, function, and nature of their GAG chains (3). GAGs are long polysaccharides formed by specific repeating disaccharide units. According to these units, GAGs can be classified as hyaluronan (HA), heparin/heparan sulfate (HS), chondroitin sulfate (CS), dermatan sulfate (DS) or keratan sulfate (KS) (4).

HS and CS/DS are highly sulfated linear GAGs composed by repeating *N*-acetylglucosamine (GlcNAc) and glucuronic acid (GlcA) disaccharides, and repeating *N*-acetylgalactosamine (GalNAc) and GlcA disaccharides, respectively. Interestingly, HS and CS/DS biosynthetic pathways, which take place at the endoplasmic reticulum-Golgi interface and in the Golgi apparatus, appear to be regulated in a coordinated way. The biosynthesis of these GAGs starts with the formation of a tetrasaccharide linker, assembled by one xylose (Xyl), two galactose (Gal) and one GlcA residue, through which both types of chains are covalently attached to their respective core proteins. This is followed by the initiation and elongation stage, during which HS- or CS-specific glycosyltransferases will dictate the synthesis of either HS or CS/DS, respectively, and catalyze the polymerization of either one of these chains. Finally, GAGs are subjected to several modification reactions mediated by different HS or CS/DS modifying enzymes and HS post-synthesis enzymes. These include epimerization of GlcA into iduronic acid (IdoA) residues and *N*-sulfation (specific for HS), *O*-sulfation and *O*-desulfation (specific for HS) reactions that confer high structural variability and functional diversity to these GAGs (46, 47).

Over the last two decades PGs and GAGs have emerged has relevant players in the field of oncobiology. Abnormal expression of numerous HS biosynthesis enzymes and HSPGs has been reported in a myriad of malignant pathologies, associating with tumor cell transformation, disease onset and progression, along with unfavorable patient outcomes (5, 347, 349). Particularly within the context of gastric cancer, recent data suggests that many HS modification and post-synthesis editing enzymes are deregulated in this pathology. Upregulation of the 6-*O*-sulfotransferase 2 (*HS6ST2)* (331) and 6-*O*-endosulfatase 2 (*HSULF2*) (335) has been reported in this disease, the former one being correlated with patients worse prognosis (331). Together, these data suggest that HS 6-*O*-sulfation motifs play an important role in this pathology. Other HS modification enzymes, like *N*-deacetylase/*N*-sulfotransferase 1 (*NDST1*) and the 3-*O*-sulfotransferase 2 (*HS3ST2*), that regulate HS *N*-sulfation and 3-*O*-sulfation respectively, have also been shown to be abnormally expressed in gastric cancer and to play a suppressive role in the disease development (330, 332).

Recently, our team has also reported upregulation of the cell surface HSPG Syndecan-4 (SDC4) on intestinal subtype gastric tumors, associated with cancer aggressiveness and patient worse prognosis, and *in vitro* cell migration and invasion capabilities (270). In accordance, in a different work we showed that glycoengineered gastric cancer cell models harboring overall increased HS content, with decreased 6-*O*-sulfation motifs, and overexpressing SDC4 presented a more aggressive malignant phenotype (388)

Interestingly, recent reports have also revealed the significant impact of CS-related enzymes and CSPGs in gastric cancer. CS synthases CHSY1, CHPF and CHSY3, three important glycosyltransferases that catalyze CS chains polymerization, were revealed to be upregulated in gastric cancer and to associate with advanced tumor stage and patients' worse prognosis. Moreover, *in vitro* and *in vivo* data showed that these enzymes play critical roles in promoting tumor cell proliferation and migration, and tumor growth (395-397). Overexpression of the extracellular CS/DS carrier biglycan in gastric cancer was also correlated with disease relapse and poor patient prognosis in advanced stages of disease (329). Additionally, CD44, particularly the CD44 variant 6 (CD44v6), a PG that can be modified with HS and CS, but also *N*- and *O*-glycans, was shown to be expressed *de novo* in stomach pre-malignant and malignant lesions (398). This glycoconjugate was also demonstrated to induce ECM remodeling, promote cell proliferation and tumor growth, and reduce cell–cell adhesions (399). Depending on its glycosylation status, CD44 was also revealed to promote activation of the receptor tyrosine kinase (RTK) recepteur d'origine nantais (RON), which is often associated with tumor cell invasion (400).

Despite the current knowledge on the variable expression of different HSPGs and GAG biosynthetic machinery in cancer, it is crucial to further learn how these changes influence

tumor cell and tissue GAG content and fine structures. Various studies have attempted to characterize the tissue GAGs expression profile and its predictive value in gastric cancer, however discrepancies between the reported data have hampered this goal (208, 324, 325).

In this work, to further uncover distinctive GAGosylation profiles in gastric cancer, we have evaluated the transcript levels of HSPGs and HS- and CS-related enzymes in a set of gastric tumor cell lines. Moreover, cellular HSPG GAGosylation profiles were also assessed to gain insight into their influence on tumor cell biology. Finally, we have performed GAG structural analysis of gastric cancer clinical samples and bioinformatic analysis to identify specific tumor-associated HS and CS/DS profiles.

RESULTS

GAG biosynthesis enzymes profiling in gastric cancer cell lines

To address the impact and roles of HS and CS/DS in the context of gastric cancer, we screened six gastric cancer cell lines derived from different histopathological backgrounds (MKN74, NCI-N87, MKN7, AGS, MKN45 and MKN1) and characterized the predominant GAGosylation profile of each cell line.

Initially, we evaluated the transcript levels of HS and CS/DS biosynthesis and post-synthesis modification enzymes (Fig. 1 and Fig. 2). In general, the Exostosin-Like 2 (*EXTL2*) and *EXTL3*, *N*-Deacetylase/*N*-Sulfotransferase 1 (*NDST1*), D-glucuronyl C5-epimerase (*GLCE*), HS 2-*O*-sulfotransferase (*HS2ST1*), HS 6-*O*-sulfotransferase 1 (*HS6ST1*) and the extracellular Sulfatase 2 (*HSULF2*) were the most frequently expressed HS-related enzymes amongst all cell lines. Regarding each enzyme family, the glycosyltransferases *EXTL2* and *EXTL3*, that catalyze main regulatory steps in HS initiation, were expressed in all cell lines (Fig. 1). Regarding HS modification reactions, *NDST1* and *NDST2* were the main NDST isoforms expressed in all gastric cancer cells, while *NDST3* was only detected in AGS and NCI-N87 cells, at residual levels in the latter, and *NDST4* was not detected in any of the analyzed cells. *GLCE* and *HS2ST1* were expressed in all cell lines. Out of the three HS6STs isoforms (*HS6ST* 1-3), that together with the HS2ST1 catalyze the most common steps of HS sulfation, *HS6ST1* was the most frequently expressed. MKN45 expressed considerable levels of *HS6ST2*, while *HS6ST3* was residually expressed in MKN1 cells (Fig. 1). Amongst the HS 3-*O*-sulfotransferases (*HS3ST* 1-6) catalyzing the rarest modification of HS, we could detect five isoforms (Fig. 1): *HS3ST1* expression was detected in 5 out of the 6 tested cell lines, *HS3ST3A* expression was detected in NCI-N87 cells, residual expression of *HS3ST3B* was detected in MKN45 and MKN1 cells, *HS3ST4* was detected in MKN1, and residual expression of *HS3ST5* was detected in all cell lines. Regarding the HSULFs 1-2, which participate in the post-synthesis editing of HS, H*SULF1*

was detected in low levels in NCI-N87, MKN7 and MKN1 cells, while *HSULF2* was expressed in all cell lines.

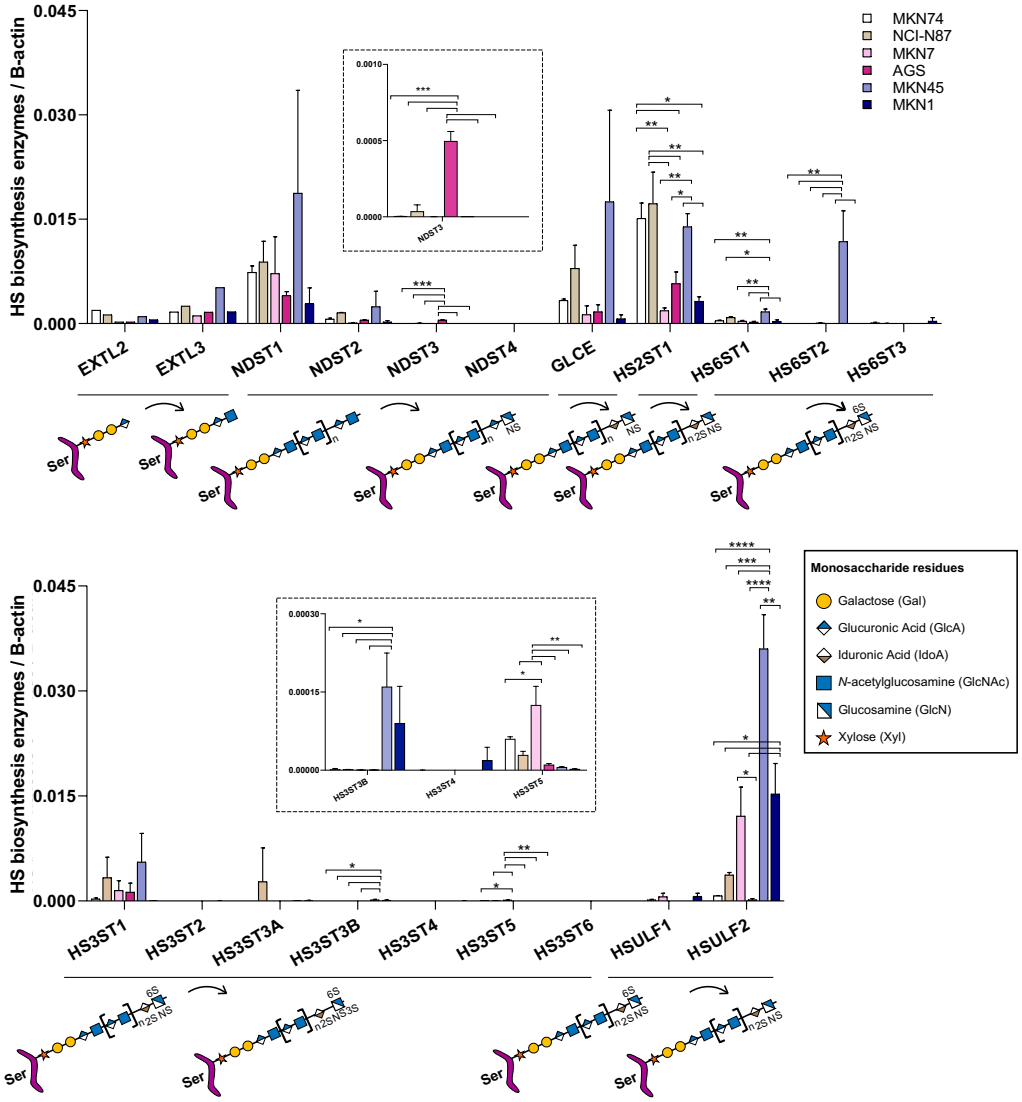


Figure 1. Transcriptomic analysis of the genes involved in HS biosynthesis and post-synthesis modifications in gastric cancer cells. The expression of the enzymes involved in HS biosynthesis (EXTL2 and EXTL3), modification (NDSTs 1-4, GLCE, HS2ST1, HS6STs 1-3, HS3STs 1-6) and editing (HSULFs 1-2) was determined for six gastric cancer cell lines *via* qRT-PCR. Bar graphs show the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. At least two independent biological assays with technical triplicates were analyzed (EXTL2 and EXTL3 were analyzed once with technical triplicates). Illustrative images depicting the different steps of HS biosynthesis catalyzed by each family of enzymes were included bellow each graph. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

Regarding CS/DS biosynthesis enzymes, CS *N*-acetylgalactosaminyltransferase 2 (*CSGALNACT2*), CS Synthase 1 (*CHSY1*), CS synthase 2 (*CHPF*), Carbohydrate Sulfotransferase 12 (*CHST12*), *CHST14*, DS Epimerase (*DSE*) and Uronyl 2-Sulfotransferase (*UST*) were expressed in all cell lines. The enzymes that initiate CS/DS synthesis, CSGALNACT1 and CSGALNACT2 presented different expression profiles between cell lines. *CSGALNACT1* was detected in most cells, except AGS, while

CSGALNACT2 was expressed in all cells (Fig. 2). Concerning the enzymes that catalyze the polymerization of CS/DS, *CHSY1* and *CHPF* were expressed in all cell lines, while *CHSY3* was only detected in NCI-N87, MKN7 and MKN1, suggesting that the remaining cell lines can conduct this particular step in the absence of this enzyme. We also evaluated the expression of the enzymes involved in CS/DS sulfation and epimerization. The expression of the 4-*O*-sulfotransferases CHST11, CHST12 and CHST13 varied between the cell lines. Unlike *CHST12*, which was expressed in all cell lines, *CHST11* was absent in AGS, and *CHST13* was only detected in MKN45 and MKN1, and at residual levels by MKN7. CHST14 (D4ST), whose 4-O-sulfotransferase activity is directed towards DS chains, was also expressed by all different cell lines but at significantly higher levels by MKN45 and MKN1. For 6-*O*-sulfotransferases, *CHST3* was mostly expressed by MKN45, MKN1, NCI-N87 and MKN7 cells, while *CHST7* was only detected in MKN45, at residual levels. Interestingly, our results also showed that *CHST15*, the 6-O-sulfotransferase that targets specifically 4-*O*-sulfated GalNAc residues, was only expressed by MKN45 cells. Both the epimerase DSE and the 2-*O*-sulfotransferase UST were expressed in all cell lines, at varying levels. Overall, we observed that at least 1 isoform required to catalyze each biosynthesis step of HS and CS is expressed in all cell lines, supporting the importance of these biosynthetic pathways in gastric cancer context.

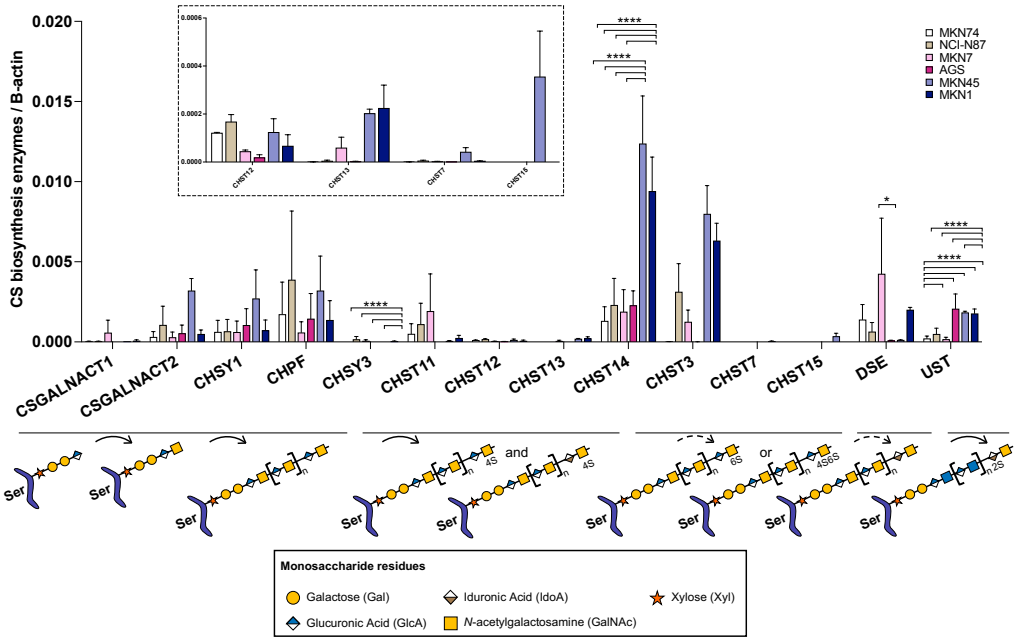


Figure 2. Transcriptomic analysis of the genes involved in CS/DS biosynthesis in gastric cancer cells. The expression of the enzymes involved in CS/DS polymerization (CSGALNACT 1-2, CHSY1, CHSY3, and CHPF) and modification (CHST3, CHST7 and CHST 11-15, DSE and UST) was determined for six gastric cancer cell lines *via* qRT-PCR. Bar graphs show the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. At least two independent biological assays with technical triplicates were analyzed. Illustrative images depicting the different steps of HS biosynthesis catalyzed by each family of enzymes were included bellow the graph. Glycan structures are represented according to the Symbol Nomenclature for Glycans (SNFG) format (1).

HSPGs profiling in gastric cancer cell lines

The majority of PGs that have been previously described in the literature to play a role in gastric cancer belong to the family of cell surface PGs, formed mostly by HS-modified proteins. Therefore, we then examined the expression of main HS carriers in epithelial cells' glycocalyx, including the transmembrane SDCs 1-4 and Glypicans 1-6 (GPCs 1-6) (Fig. 3). We observed that only *SDC1*, *SDC4* and *GPC1* were expressed in all of the six analyzed gastric cancer cell lines. *SDC2* was expressed more abundantly by MKN1 cells, followed by MKN7 cells, and *SDC3* was expressed at very low levels by all cell lines, except for MKN74 that did not express this PG. *SDC4* was greatly expressed by MKN45 cells, followed by NCI-N87 cells, which presented higher *SDC4* transcript levels than MKN74, MKN7 and AGS cells (Fig. 3). Regarding GPCs, *GPC1* expression was higher in NCI-N87, followed by MKN1; *GPC3* was only expressed by MKN45 cells; and *GPC4* was significantly more expressed by MKN45 cells. *GPC2*, *GPC5* and *GPC6* were not detected in any of the gastric cancer cell lines analyzed (Fig. 3).

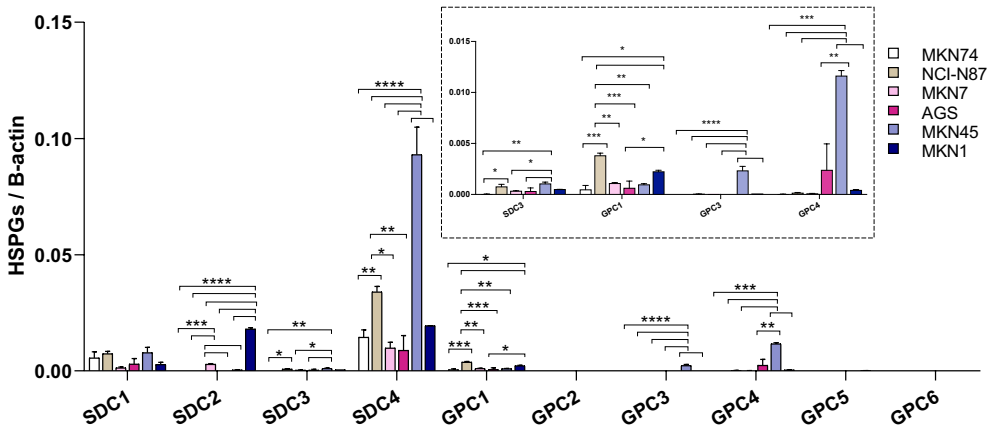


Figure 3. Transcriptomic analysis of cell surface HSPG genes expressed in gastric cancer cells. The expression of HSPGs, Syndecans (SDCs) 1-4 and Glypicans (GPCs) 1-6 was determined for six gastric cancer cell lines *via* qRT-PCR. Bar graph shows the mean RQ values for each gene, normalized to the internal control gene β-actin + SD. At least two independent biological assays with technical triplicates were analyzed.

Addressing gastric cancer cell HS/CS GAGosylation features

To characterize the HS expression profile in each gastric cancer cell line, WB analysis was performed to measure total cellular levels. We observed the HS specific smeared staining by 10E4 at higher molecular weight (MW), ranging between 150-250 kDa in all cell lines, which was fully abolished upon HS digestion by Heparinase III (Hep. III). This staining was more intense in MKN7, MKN45 and MKN1 cells, suggesting higher levels of HS in these particular cell lines (Fig. 4A).

To further infer about HS content, we also evaluated the number of HS chains in each cell line by digesting the protein lysates with Hep. III followed by 3G10 immunoblotting. Notably, 3G10 staining of the HS stubs followed a different trend in terms of overall intensity

(Fig. 4B). MKN74 cells, which presented lower levels of HS content compared to MKN7, MKN45 and MKN1 cells (Fig. 4A), exhibited similar high 3G10 signal to MKN7 and MKN1 cells (more bands with higher intensity; Fig. 4B), suggesting the presence of a higher number of HS chains, that might be shorter. On the other hand, MKN45 cells, showing higher levels of HS (Fig. 4A), exhibited low signal of HS stubs, suggesting the presence of fewer but longer HS chains. NCI-N87 and AGS cells presented similar 3G10 labeling profiles (Fig. 4B). Overall, these differences support a varying repertoire of HSPGs, labeled at different MW, among the different gastric cancer cell lines.

We then proceeded with the evaluation of CS levels. Protein lysates were treated with or without Hep. III and/or Chondroitinase ABC (Chond. ABC), to enhance the antibody binding, as advised by the antibodies manufacturer, and immunoblotted for Chondroitin-4-sulfate (C4S) and Chondroitin-6-sulfate (C6S) motifs using the BE-123 mAb and MK-302 mAb, respectively (Fig. 4C and D). C4S labeling varied greatly between each cell line, as NCI-N87, MKN7 and MKN1 cells presented higher labeling intensity at higher MW (Fig. 4C). Similar to what we observed for 3G10, the labeling profile of the C4S stubs was cell line dependent, suggesting a specific CS-modified PG repertoire labeled at different MW. Interestingly, in the MKN7 cell line C4S labeling was also detected in the non-digested condition, which might indicate higher abundance of this CS motif. Noteworthy, upon either HS digestion and 3G10 labeling (Fig. 4B), or CS digestion and BE-123 labeling (Fig. 4C), NCI-N87 labeling profiles revealed a similar strong double band close to 37 kDa, suggesting that these cells synthesize a CS and HS bearing core protein, possibly corresponding to a hybrid PG. According to its MW, this PG could be a SDC, more specifically SDC4.

Regarding C6S, the labeling profile and intensity were similar between most cell lines as these motifs were labeled as two bands above 150 and 250 kDa that were dependent on enzymatic digestion, and one band around 75 kDa that was independent from Hep. III and Chond. ABC digestion (Fig. 4D). Noteworthy, MKN45 exhibited a slightly different and more intense labeling profile (Fig. 4D).

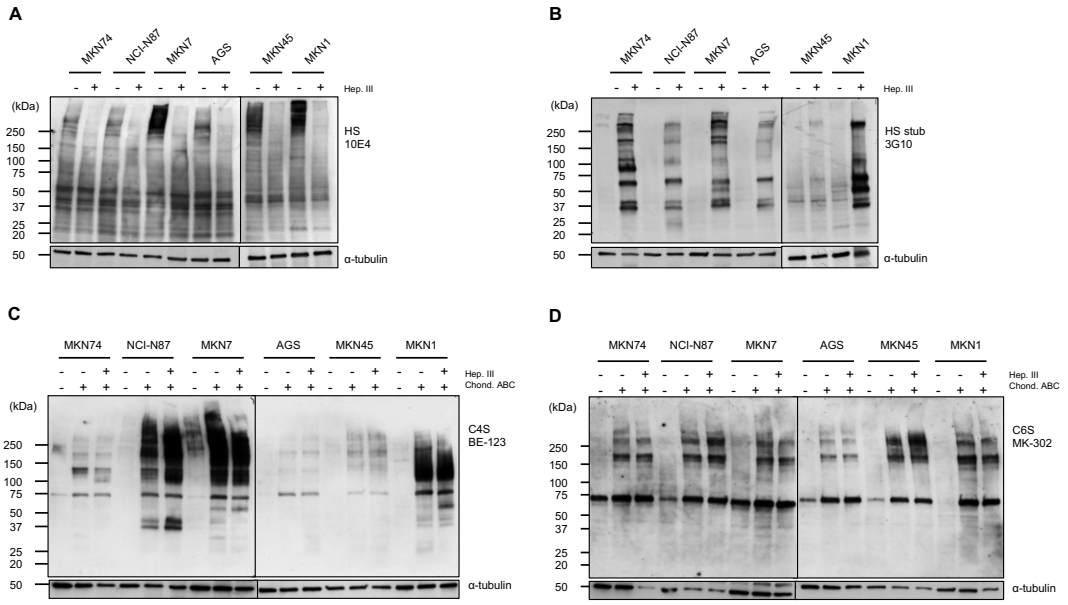


Figure 4. Evaluation of HS and CS overall quantity and chains number amongst different gastric cancer cell lines. The six gastric cancer cell lines were evaluated for HS total content and number of HS chains by WB analysis. Protein lysates were treated with or without Heparinase III (Hep. III) and immunoblotted for **(A)** undigested HS using the 10E4 mAb or **(B)** HS stubs using the 3G10 mAb. Images are representative of 2 independent biological assays. **(C and D)** CS levels were also evaluated *via* WB analysis. Protein lysates were treated with or without Chondoritinase ABC (Chond. ABC) and Heparinase III (Hep. III), to enhance the antibody labeling, as advised by the antibodies manufacturer, and immunoblotted for **(C)** Chondroitin-4-sulfate (C4S) and **(D)** Chondroitin-6-sulfate (C6S) motifs using the BE-123 mAb and MK-302 mAb, respectively.

SDC1 and SDC4 expression and GAGosylation profiles in gastric cancer cells

Taking in consideration the different HS/CS profiles observed across the different cell lines, we next investigated the protein levels and GAGosylation of the two transmembrane HSPGs expressed in all gastric cancer cell lines (Fig. 3): SDC1 and SDC4 (Fig. 5). In all cell models, SDC4 was labeled as a single band around 37 kDa (non-glycosylated form) and as a smeared staining at higher MW (glycosylated forms) (Fig. 5). WB analysis revealed lower levels of non-glycosylated SDC4 in MKN74 cells when compared with the other five cell lines, however it showed a more intense smeared staining of the glycosylated SDC4 forms. The remaining cell lines, though presenting similar levels of non-glycosylated SDC4, showed different SDC4 smeared labeling intensities. Together with MNK74, NCI-N87 and MKN45 cells exhibited higher levels of glycosylated SDC4, when compared to MKN7, AGS and MKN1 cell lines (Fig. 5). SDC4 GAGosylation was also evaluated by digesting the protein lysates with Hep. III and/or Chond. ABC. Although SDC4 labeling profile was resistant to Chond. ABC digestion in all six cell lines, double digestion further shortened the labeling smear observed in the Hep. III single digested samples, particularly in NCI-N87, MKN7 and MKN45 cells (Fig. 5). This observation suggests that SDC4 molecules in these cell lines might be modified with both HS and CS, and that Chond. ABC digestion is more efficient towards HS-naked SDC4 molecules (explaining therefore the lack of labeling

variations upon single Chond. ABC digestion). We cannot also dismiss the possibility that these differences could be due to the longer digestion periods that double digested samples are subjected to (since these samples underwent two overnight incubation periods at 37ºC: first with Hep. III, and then after addition of Chond. ABC).

Regarding SDC1, labeling with B-A38 antibody was only observed upon Hep. III digestion, indicating that HS chains modifying this HSPG might be too long and heavy and promote steric hindrance to the antibody recognition of the protein core. MKN74 and NCI-N87 cells showed higher levels of SDC1, while in MKN7, MKN45 and MKN1 cells the labeling of SDC1 was only observed when using longer ECL exposure time (Fig. 5). SDC1 was not observed at the protein level in AGS cells by immunoblot, though transcript levels were detected (Fig. 3).

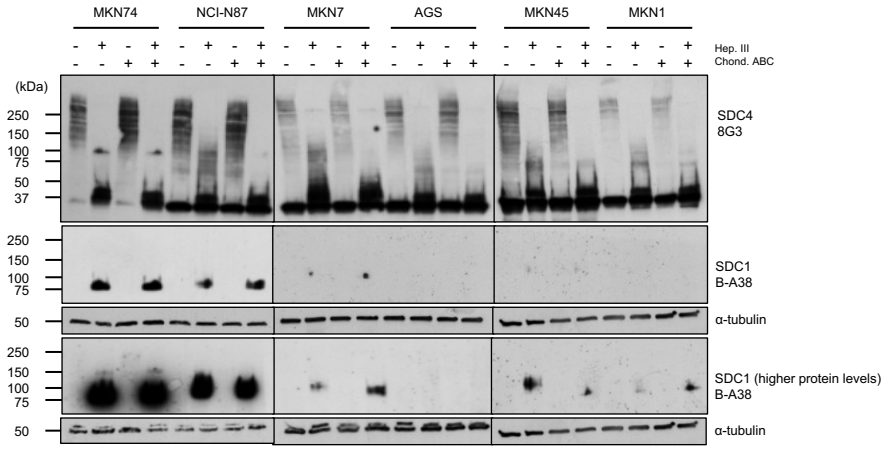


Figure 5. Evaluation of SDC1 and SDC4 levels and GAGosylation profiles in gastric cancer cells. The protein levels and GAGosylation profiles of SDC4 and SDC1 were evaluated by WB analysis. Protein lysates were treated with or without Heparinase III (Hep. III) and Chondroitinase ABC (Chond. ABC) and immunoblotted for SDC4 and SDC1 using 8G3 and B-A38 mAb, respectively. Images are representative of technical duplicates.

Identification of GAG transcriptomic signatures in gastric tumors

To further explore the expression of specific GAG-related genes in gastric cancer patients, we analyzed the transcriptomic profile of several genes using samples from a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STDA). We selected genes coding for specific PGs and GAG biosynthetic enzymes, including membrane bound HSPGs (SDCs 1-4 and GPCs 1-6) HS polymerizing enzymes (EXTLs 2-3, EXTs 1-2), HS 6-*O*-sulfation regulatory enzymes (HS6STs 1-3 and HSULFs 1-2), and CS polymerizing enzymes (CSGALNACTs 1-2, CHSY1 and 3, CHPFs 1-2).

TCGA data mining revealed that the expression of the HSPGs *SDC4* and *GPC2* is higher in intestinal subtype gastric cancer compared to healthy individuals, while *GPC6* expression was found to be downregulated (Fig. 6A). Interestingly, regarding HS initiation and elongation enzymes, the HS negative regulator *EXTL2* is significantly upregulated in

gastric cancer diffuse subtype, while the HS polymerizing enzymes *EXT1* and *EXT2* are significantly upregulated in intestinal and diffuse gastric cancer (Fig. 6B).

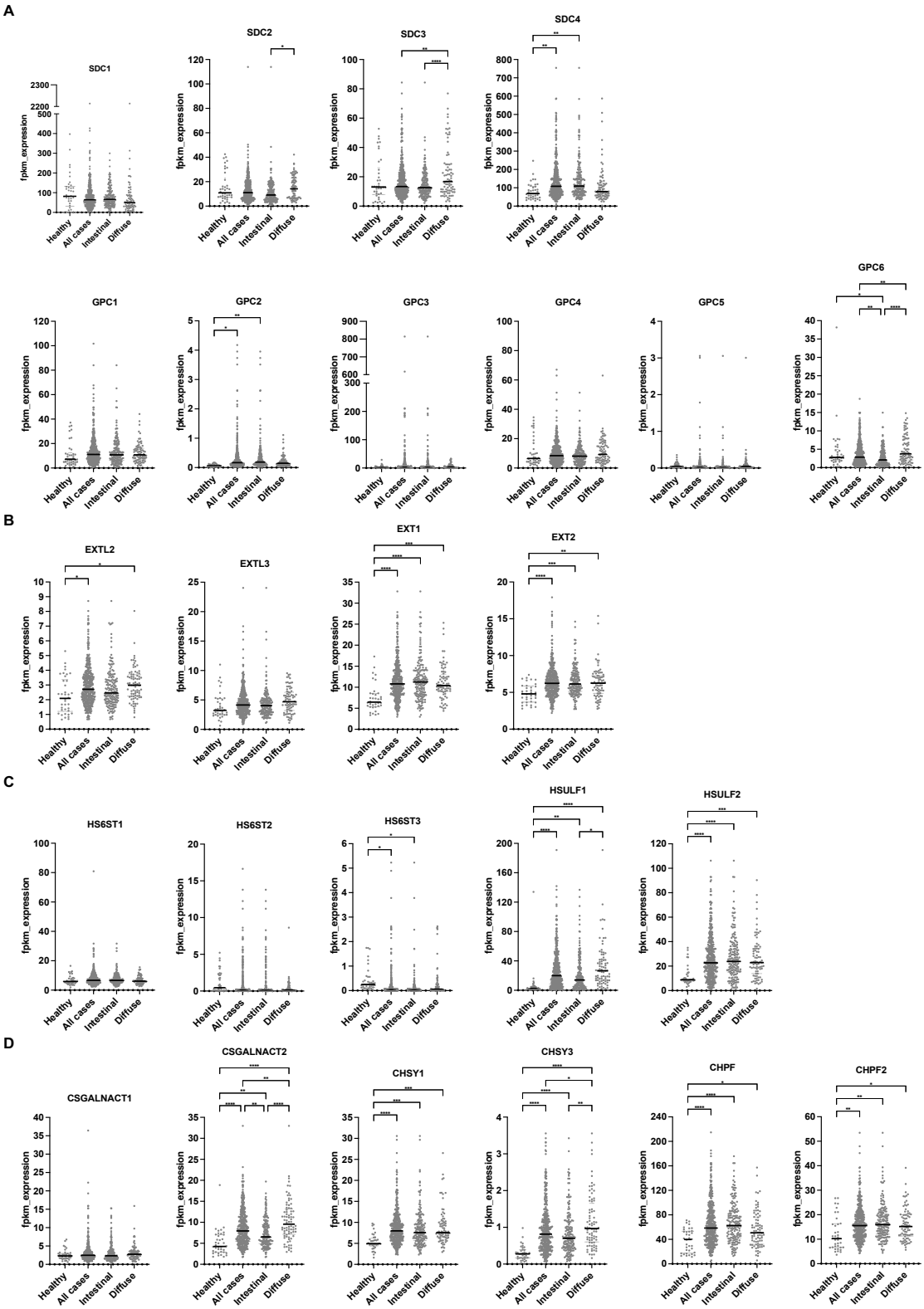


Figure 6. Identification of a GAG-related gene expression profile in gastric cancer. Distribution of the levels of mRNA of **(A)** HSPGs, **(B)** HS glycosyltransferases**, (C)** HS modification and post-synthesis editing and **(D)** CS glycosyltransferases in a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD) were evaluated following to the Lauren classification. The “all cases” category includes the data from diffuse, intestinal and other gastric cancer subtypes. mRNA gene expression was quantified as FPKM (fragments per kilobase million).

Concerning the enzymes that modulate HS 6-*O*-sulfation, *HS6ST3* is significantly downregulated in intestinal gastric cancer, while both *HSULF1* and *HSULF2* are upregulated in intestinal gastric cancer (Fig. 6C), suggesting that decreased glycocalyx and extracellular HS 6-*O*-sulfation might favor oncogenic events. Finaly, TCGA data analysis strongly indicates that, with the exception of *CSGALNACT1*, the enzymes responsible for CS initiation and polymerization are significantly upregulated in gastric cancer, both intestinal and diffuse subtypes (Fig. 6D).

Identification of GAG structural signatures in gastric tumor clinical samples

To address the clinical relevance of distinct HS sulfation profiles and aiming to identify unique gastric tumor-associated GAGosylation signatures, we performed GAG disaccharide structural analysis of tumor tissue samples collected from advanced gastric cancer patients (n=5; m= <100mg per sample) paired with tissue samples from distant mucosa of the same individuals. For processing, the tissues were mechanically disrupted with a Potter grinder in denaturing (Urea) buffer. Proteins were degraded by pronase digestion and precipitated by incubation with ice-cold trichloroacetic acid (TCA), while lipids were removed through diethyl ether washes (Fig. 7A). GAGs were then purified using the same protocol used for cellular GAGs (ion-exchange chromatography and size exclusion chromatography) and analyzed *via* reversed-phase ion pairing (RPIP)-HPLC based disaccharide analysis. We were able to optimize and validate this protocol and detected significant levels of HS and CS/DS disaccharides using small quantities of tissue samples (Fig. 7B and C).

HS and CS/DS disaccharide levels and sulfation profiles were analyzed, and preliminary data from two patients are suggestive of decreased overall HS *O*-sulfation in the tumor tissue (T), compared to the normal mucosa (ND; Fig. 7B). The structural data is not suggestive of a specific trend regarding the variations in disaccharide composition, neither in HS or CS/DS (Fig. 7B and C). Nonetheless, additional analysis should be performed on a larger cohort to validate a specific disaccharide structural profile.

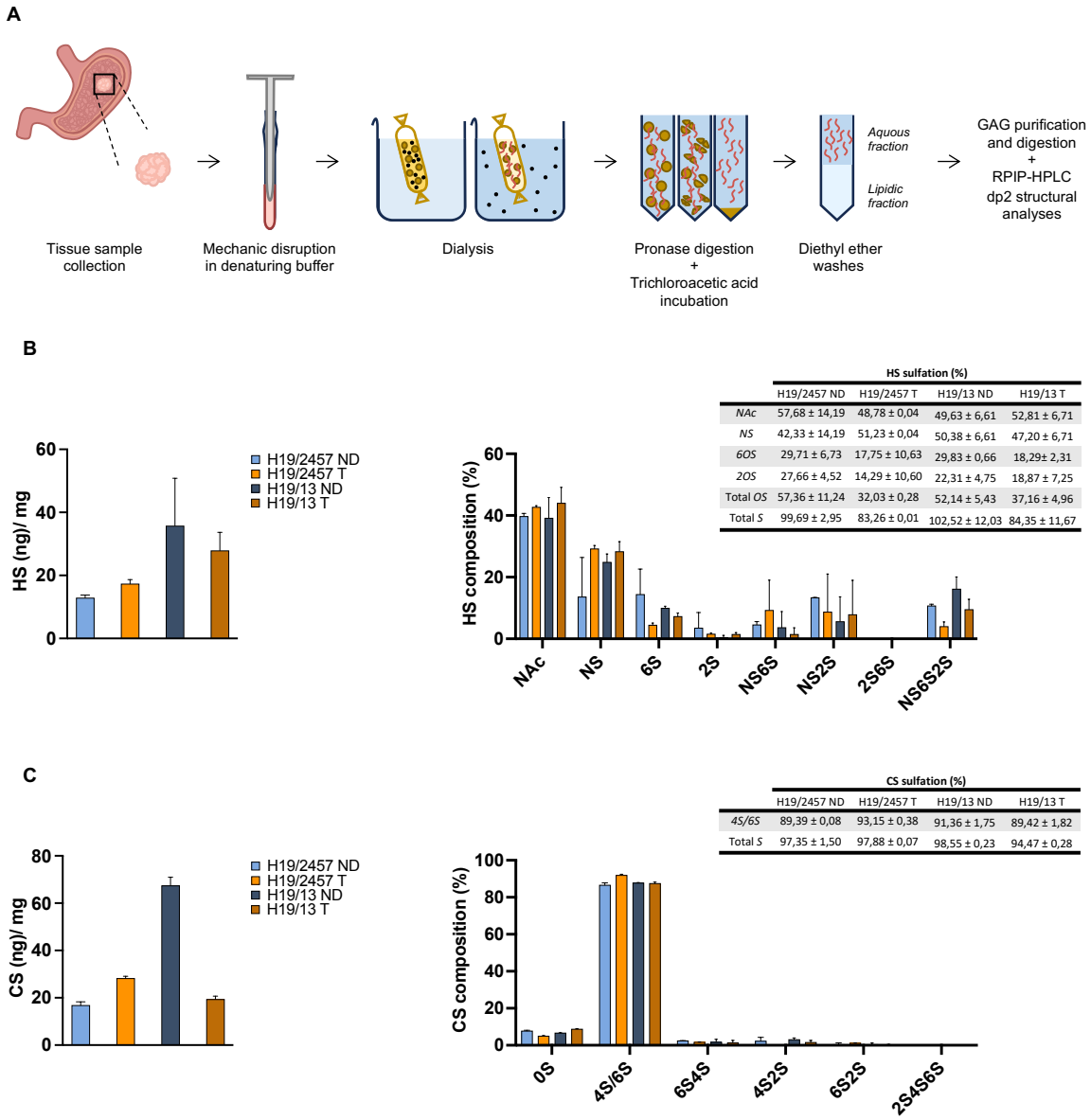


Figure 7. Identification of a GAG structural signature in gastric cancer tissues. (A) The total amounts and structural composition of HS and CS disaccharides from gastric tissue samples were analyzed. GAG chains were isolated *via* mechanic disruption, protein digestion and precipitation, and lipids removal. GAGs were then purified and digested with a mixture of Heparinases I, II and III or Chondroitinase ABC, and analyzed by RPIP-HPLC. **(B**; *right***)** Bar graph shows the mean levels of HS disaccharides, expressed in ng. Technical duplicates were analyzed. **(B**; *left***)** Relative quantities of distinct HS disaccharide units were measured. Bar graph shows the mean levels of each type of HS disaccharide: NAc = ΔHexUA-GlcNAc; NS = ΔHexUA-GlcNS; 6S = ΔHexUA-GlcNAc,6S; 2S = ΔHexUA,2S-GlcNAc; NS6S = ΔHexUA-GlcNS6S; NS2S = ΔHexUA,2S-GlcNS; 2S6S = ΔHexUA,2S-GlcNAc,6S; NS2S6S = ΔHexUA,2S-GlcNS6S. Table represents the mean values of sulfation content. NAc = *N*-Acetylation; NS = *N*-Sulfation; 6S = 6-*O*-Sulfation; 2S = 2-*O*-Sulfation; total OS = total *O*-Sulfation, total S = total Sulfation (*N*- and *O*-sulfation). **(C**; *right***)** Bar graph shows relative mean quantities of distinct CS disaccharide units. Technical duplicates were analyzed. **(C**; *left***)** Relative quantities of distinct CS disaccharide units were measured. Bar graph shows the levels of each type of CS disaccharide: 0S = ΔHexUA-GalNAc; 4S = ΔhexUA-GalNAc,4S; 6S = ΔhexUA-GalNAc,6S; 6S4S = ΔhexUA-GalNAc,6S4S; 4S2S = ΔhexUA,2S-GalNAc,4S; 6S2S = ΔhexUA,2S-GalNAc,6S; 2S4S6S = ΔhexUA,2S-GalNAc,6S4S. Table represents the CS sulfation content in each cell model. 4S/6S = 4-*O*-Sulfation and 6-*O*-Sulfation combined; total S = total Sulfation.

Addressing the impact of GAG signatures in gastric cancer outcome

To further infer on the impact of HSPGs and GAG enzymes on disease progression, Kaplan-Meier (KM) analysis of several HSPGs and all HS and CS biosynthetic enzymes was performed using the same TCGA-STAD cohort. These data are summarized in the dot plot depicted on Figure 8, representing the genes whose up- or downregulation are associated with lower disease-free survival (DFS) and lower overall survival (OS).

A large number of altered genes influencing patient outcomes were detected in intestinal subtype gastric cancer. Focusing on this specific subtype, increased expression of several membrane-bound PG coding genes appears to significantly affect patient's prognosis. *SDC1* downregulation appears to promote worse outcomes, while *SDC4* overexpression correlates with lower DFS, in accordance with our previous observations (270). Additionally, and in contrast with the expression analysis (Fig. 6A), data suggest that *GPC6* upregulation also associates with lower OS (Fig. 8). For linker enzymes, several downregulated genes, including *O*-Xylosyltransferases (*XYLTs* 1-*2*) and 2-Phosphoxylose phosphatase *(PXYLP1)*, were associated with lower DFS and OS. Similarly, decreased expression of the HS glycosyltransferases *EXTL3* and *EXT2* was linked to lower DFS.

Regarding HS modification enzymes, no specific trend was observed, as the expression of different enzymes that catalyze different steps of modification appears to influence patients' outcome differently. In general, upregulation of *NDSTs* and *HS6ST2* and downregulation of *GLCE* were associated with lower DFS and OS. Regarding HS3STs, up- and downregulation of different isoforms were associated to lower OS, suggesting that different isoforms may promote the formation of distinct 3-*O*-sulfation epitopes, with variable impact on this pathology (as observed for other pathologies (401)). Finally, the upregulation of both *HSULF1* and *HSULF2*, also described in the expression analysis (Fig. 6C), was associated with poor patient prognosis. In terms of CS/DS biosynthesis, increased expression of the polymerizing enzymes *CSGALNACT2* and *CHSY3* and decreased expression of *CHPF2* was associated with lower OS and DFS (Fig. 8). The different CS/DS glycosyltransferases (CHSY1, CHSY3, CHPF and CHPF2) are known to pair up to catalyze chains elongation and according to the literature, polymerization takes place even in the absence of some of these enzymes (356). Nonetheless, the regulation of this specific step is still largely unknown, as well as the potential outcomes in CS/DS synthesis resulting from the coupling of different enzymes. This might justify the opposite trends in the expression of the above-mentioned genes and how it affects patients' disease. Finaly, CS/DS modification reactions feature a marked trend, with the overexpression of genes encoding different CS sulfation and epimerization enzymes being associated with lower DFS and OS (Fig. 8).

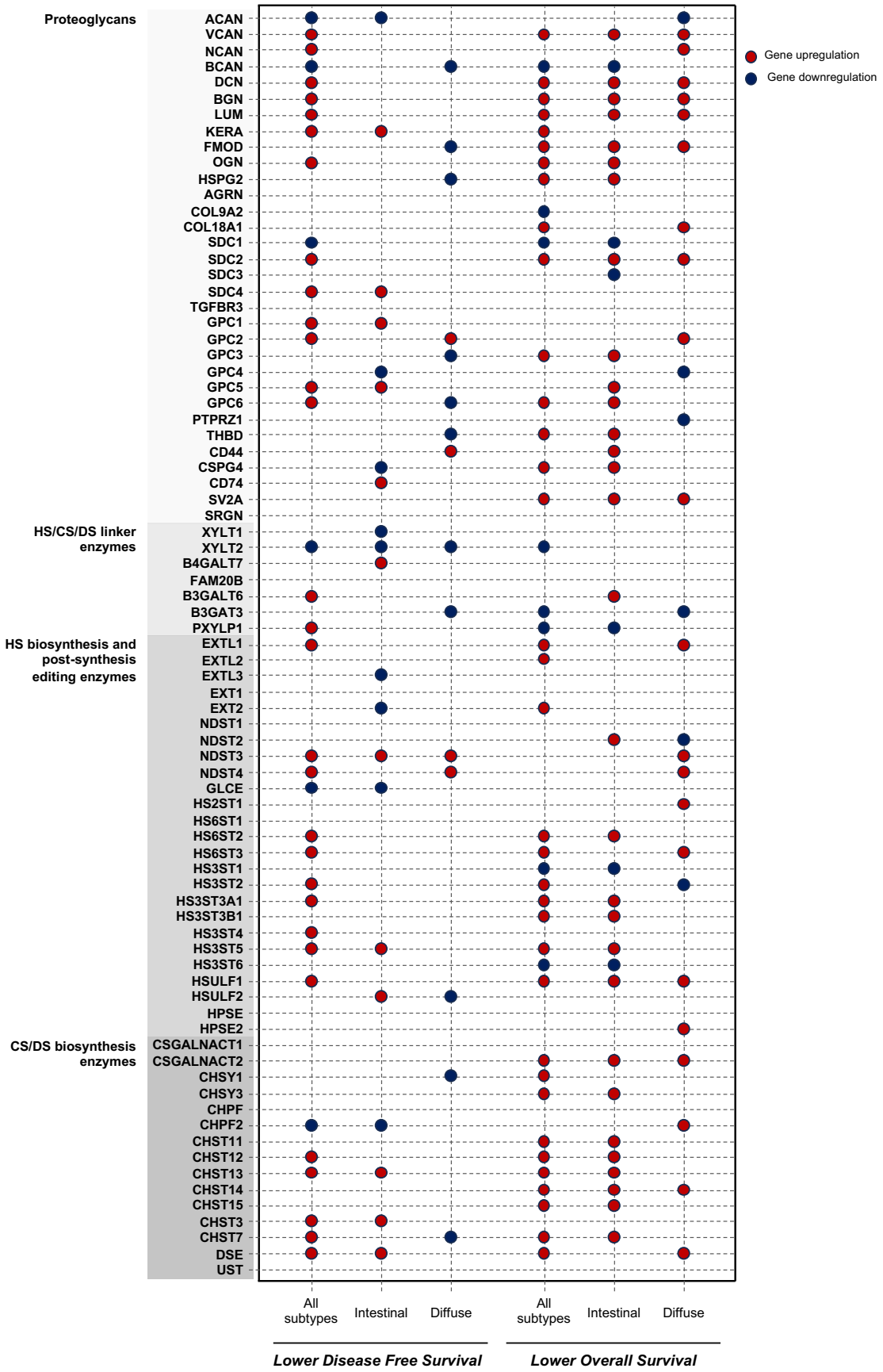


Figure 7. Addressing the impact of GAG-related gene expression patterns in gastric cancer patient survival outcomes. Dot plot graphic representation of the genes encoding several PGs and all HS and CS/DS biosynthetic enzymes, whose up- (*red dot*) or downregulation (*blue dot*) associated with lower disease-free survival (DFS) and lower overall survival (OS). DFS and OS data was assessed *via* Kaplan-Meier (KM) analysis using a gastric cancer patient cohort from The Cancer Genome Atlas Stomach Adenocarcinoma (TCGA-STAD).

For the diffuse subtype, fewer membrane bound HSPGs and HS polymerizing enzymes were associated with OS (only *EXTL1* upregulation was associated with lower OS). On the other hand, and similar to what we observed for the intestinal subtype, downregulation of linker enzymes coding genes (*XYLT2 and B3GAT3*) was associated with lower DFS and OS. No specific trend regarding the expression of HS modifying enzymes correlated to lower DFS or OS (Fig. 8). Regarding CS/DS biosynthesis, the expression of fewer genes was shown to influence the disease outcome: downregulation of *CHSY1* and CHST7 was associated with lower DFS, while upregulation of *CSGALNACT2*, *CHPF2*, *CHST14* and *DSE* was correlated with lower *OS* (Fig. 8).

DISCUSSION

Many studies over the years have established the important functions displayed by PGs and GAGs in cell malignant transformation, ranging from modulation of oncogenic cell signaling, cell motility and invasion capabilities and tumor cell-microenvironment cross-talk, to tumor growth, metastasis and angiogenesis (94, 402). The identification of specific cancer associated HS and/or CS motifs, either in terms of total amount or distinct sulfation patterns, might lead to the discovery of new glyco-based biomarkers with potential for early cancer detection and patients' stratification.

To our knowledge, this study is the first profiling of HS- and CS/DS-related genes and corresponding cellular GAGosylation features in a set of gastric cancer cell lines originated from different histopathological gastric cancer settings. In our study, qRT-PCR analysis showed that six different gastric cancer cell lines express the essential enzymatic machinery to carry all major steps of HS and CS/DS biosynthesis. At least one enzyme isoform of each class of enzymes was detected in all cell lines, except for the class of CS 6-*O*-sulfotransferases (*CHST3, 7* and *15*) that were not expressed by MKN74 and AGS cells (Fig. 1 and 2). Regarding HS biosynthesis, aligned with the literature, *NDST1* and *NDST2*, which are the main NDSTs isoforms expressed in all vertebrates' adult tissues, were detected in all six cell lines (Fig. 1). These enzymes catalyze the first HS modification step, on top of which most of the remaining modification reactions are conducted (epimerization and *O*-sulfation). Additionally, different NDST isoforms are endowed with different *N*-deacetylase and *N*-sulfotransferase activity ratios, therefore the variable expression of these enzymes could account for the possible diverse HS sulfation profiles (351, 403). Both *GLCE* and *HS2ST1* exhibited a similar expression trend between the different cell lines (Fig. 1). This is quite interesting since the activity of HS2ST1 appears to rely on the activity of the epimerase, as it shows higher affinity towards epimerized IdoA, compared to GlcA residues (159). *HS6ST1* was the main isoform expressed amongst all cell lines, despite HS6STs having broad substrate recognition and enzymatic activity (Fig. 1). Interestingly, *HS3STs*

expression, although low, varied considerably between the tested cell lines (Fig. 1). Knowing that different HS3ST isoforms have different substrate specificities (401, 404), we can speculate that these differences might contribute to establish different HS 3-*O*-sulfate epitopes, with distinct roles in tumor cells. Finally, *HSULF2,* commonly described as being pro-oncogenic, was detected in all cell lines, while *HSULF1* had a more restricted expression. Noteworthy, AGS cells that expressed the least amount of *HS6STs,* did not express *HSULF1* and expressed the lowest levels of *HSULF2*, hinting that HS *6-O*-sulfation might be downregulated in this cell line (Fig. 1). In addition, it alludes to a possible concerted self-regulatory mechanism for the expression of HS-associated genes participating in correlated modification reactions, which has been previously theorized (405).

Regarding CS/DS biosynthesis, the expression of initiation and polymerization enzymes varied comparing all cell lines, with *CSGALNACT2*, *CHSY1* and *CHPF* being the predominant isoforms (Fig. 2). This suggests that in certain cell lines, these enzymes can catalyze CS elongation alone, which was further confirmed by WB analysis, through which we were able to detect C4S in all gastric cancer cell lines, although at different levels (Fig. 4C). Interestingly, C4S staining was lower in AGS and MKN45 cells (Fig. 4C), which presented considerable expression of the CS/DS polymerizing enzymes (Fig. 2). Nonetheless, it is likely that these cells harbor larger levels of CS that is not modified with 4-*O*-sulfated (or 6-*O*-sulfated) motifs, and therefore were not detected in our WB analysis. CS 4-*O*-sulfotransferases (CHST11-13) were detected at low levels in different cell lines, while CS 6-*O*-sulfotransferases (CHST3, 7 and 15) were the only evaluated enzymes that were not expressed by all cell lines, as we could not detect any of the isoforms in MKN74 and AGS. These data are in accordance with the CS structural data previously reported on MKN74 cells, in which C6S disaccharides were not detected (388). Overall, the reduced or absent expression of CS sulfotransferase enzymes in the gastric cancer cell lines is consistent with previous studies reporting a significant increase of non-sulfated CS/DS content in gastric cancer samples (325). Interestingly, DS specific enzymes, CHST14 and DSE were expressed in all cell lines (Fig. 2).

The function of membrane bound PGs in cell malignant transformation and tumor progression has been extensively explored over recent years. Particularly in the context of gastric cancer, abnormal expression of HSPGs like SDC1 (326), SDC4 (270) and GPC6 (328) has been reported. In our study, we were able to detect HS, *SDC1*, *SDC4* and *GPC1* in all gastric cancer cell lines (Fig. 3-5). Although with different labeling intensities, the SDC4 staining smear covered a similar MW range in all cell lines, suggesting that the length of SCD4 HS chains did not vary (Fig. 5). Noteworthy, HS and CS labeling data suggested the existence of a hybrid SDC molecule, whose GAGosylation profile appears to be gastric tumor cell dependent (Fig. 4 and 6). We have previously reported that SDC4 is modified

with CS chains in MKN74 gastric cancer cells lacking EXTL3 and therefore harboring impaired HS biosynthesis (388). Additionally, we observed that in NCI-N87 cells, hybrid SDC4 might also occur despite functioning HS biosynthesis.

Attempting to identify new therapeutic targets, several studies have explored the expression of GAGs and GAG related genes in various malignancies. GAG disaccharide structural analysis revealed increased in CS content and decreased CS sulfation in lung tumors, while HS did not vary but exhibited increased *O*-sulfation motifs (406). On the other hand, in hepatocellular carcinoma, a different trend showing increased HS content and reduced HS 6-*O*-sulfation was described (310). Transcriptomic analysis also revealed gene deregulation of several HS and CS biosynthesis enzymes and HSPGs in breast (209) and colorectal (286, 287) cancer, a few of which were dependent on metastatic potential. Similar analysis also revealed general downregulation of HS-biosynthesis associated transcriptome in prostate cancer (407). In our study, TCGA data mining revealed interesting expression trends of GAG-related genes in gastric cancer. Generally, we confirmed the previously reported pattern of *SDC4* overexpression in intestinal gastric cancer (270). However, *GPC6* was downregulated (Fig. 6A), in opposition to previous studies describing increased *GPC6* expression in intestinal gastric tumors (328). Interestingly, TCGA data further showed a significant upregulation of the biosynthetic machinery catalyzing HS and CS polymerization in both diffuse and intestinal subtypes (Fig. 6B and D), although the negative regulator of HS biosynthesis, *EXTL2*, also appeared to be augmented, but only in the diffuse subtype (Fig. 6B).

HS 6-*O*-sulfotransferases and sulfatases regulate HS 6-*O*-sulfation levels in different manners: HS6STs promote overall HS 6-*O*-sulfation, while HSULFs remove 6-O-sulfate groups from HS trisulfated disaccharides (IdoA2SGlcNS6S) in highly sulfated domains, and in a processive manner (186). The expression profile of HS6STs and HSULFs suggests a potential reduction of HS 6-*O*-sulfation motifs in gastric cancer, as *HS6ST3* is downregulated in intestinal subtype carcinomas, while *HSULF1* and *HSULF2* are upregulated in both the intestinal and diffuse subtypes (Fig. 6C). These data contrast with the general trend reported for different malignancies. *HS6ST3* increased expression has been reported in cartilage (297) and breast (299) tumors, together with increased 6-*O*-sulfated HS disaccharides in high-grade chondrosarcoma cell lines (297) and in the ovarian endothelium of patients with advanced stage ovarian cancer (300). It will be important to further dissect the role that HS 6-*O*-sulfation plays in gastric cancer and which signaling cascades might be influenced by this specific structural motif.

Overall, by screening six gastric cancer cell lines, we revealed gastric tumor cell specific GAG transcription profiles and HSPG GAGosylation features. Our results underline the high variability of these structural elements, even within different cell types of the same

malignancy, which might pose a challenge to our initial goal of identifying a cancer specific GAG signature. Considering this goal, we further optimized and employed (RPIP)-HPLC based tissue GAG disaccharide analysis, and preliminary data supported that this methodology provides high resolution separation and sensitive detection of HS and CS/DS structural features using small tissue samples (Fig. 7).

In addition, to validate the clinical relevance of HS and/or CS/DS profiles in cancer, we analyzed samples from a gastric cancer patient cohort from TCGA-STAD and observed an interesting transcriptomic trend suggesting overall increased HS and CS formation, and decreased HS 6-*O*-sulfation in gastric malignancy. While, it was not possible, at this stage, to pinpoint the specific influence of these changes on patient's prognosis, an overall upregulation of the genes encoding for CS modification enzymes correlated with patients lower DFS and OS. Although further studies are required in order to define a gastric cancer specific GAGosylation profile, our data revealed a transcriptomic signature of the GAG biosynthetic machinery supporting the deregulation of the HS/CS metabolic system in gastric cancer.

EXPERIMENTAL PROCEDURES

Cell lines and cell culture

Six gastric cancer cell lines were used in this study: MKN74 (JCRB0255), MKN45 (JCRB0254), MKN7 (JCRB1025) and MKN1 (JCRB0252), obtained from the Japanese Collection of Research Bioresources Cell Bank, and AGS (CRL-1739™) and NCI-N87 (CRL-5822™), acquired from American Type Culture Collection (ATCC). Cells were cultured in RPMI Medium 1640 (1X) + GlutaMAX™-I (Gibco), supplemented with 10% (v/v) Fetal Bovine Serum (FBS) (Biowest) and maintained at 37°C under 5% (v/v) CO_2 conditions.

Antibodies

The antibodies used in this manuscript and the conditions used for the analysis described below are listed in Table 1.

Table 1. *List of antibodies and conditions used for Western Blot (WB).*

Antibodies	Dilutions	Source/Reference
HS (F58-10E4)	1:600	Amsbio
HS stub region (F69-3G10)	1:500	Amsbio
C4S (BE-123)	1:1000	Sigma-Aldrich
C6S (MK-302)	1:1000	Sigma-Aldrich
SDC1 (B-A38)	1:125	Abcam
SDC4 (8G3)	1:2000	(386)
α-Tubulin (DM1A)	1:10000	Sigma-Aldrich
Horseradish Peroxidase-conjugated α mouse IgG	1:5000	Jackson IR Laboratories, Inc
Horseradish Peroxidase-conjugated α mouse IgG1	1:8000	Jackson IR Laboratories, Inc
Horseradish Peroxidase-conjugated α mouse IgM	1:5000	Jackson IR Laboratories, Inc

Western Blotting

Total protein lysates were obtained from confluent cells scraped and lysed with Lysis Buffer 17 (R&D Systems) supplemented with cOmplete™ (Roche) and PhosSTOP™ (Roche) inhibitors cocktail. The protein concentration of each lysate was quantified using the DC Protein Assay (BioRad). Total protein lysates (12,5 µg for HS stub analysis and 25 µg for HS, SDC1 and SDC4, C4S and C6S analysis) were incubated with 4x Laemmli Sample Buffer (BioRad) supplemented with 10% β-mercaptoethanol, at 98°C for 5 min. The protein extracts were loaded and run in 4-15% gradient precast polyacrylamide gels (BioRad) and then transferred onto a nitrocellulose membrane (GE Healthcare Life Sciences). The membranes were blocked for 1h at room temperature with different solutions depending on the primary antibodies applied next: TBS 5% BSA and 0.05% Tween® 20 (TBS-T 0.05%; Sigma-Aldrich) buffer, PBS 5% BSA and 0.05% Tween® 20 (PBS-T 0.05%) buffer, PBS-T 5% non-fat dry milk or TBS-T 5% non-fat dry milk for 3G10, 10E4, 8G3 or B-A38, respectively (Table 1). The membranes were incubated overnight at 4°C with primary antibodies (Table 1) and then washed with TBS-T 0.05% or PBS-T 0.05%, followed by incubation for 1h at room temperature with horseradish peroxidase-conjugated secondary antibodies. Each antibody used in this assay was diluted in the respective blocking solution. For the membrane's development, new washes with TBS-T 0.05% or PBS-T 0.05% were performed, and the protein bands visualized with ECL chemiluminescent western blotting detection reagent and films (GE Healthcare Life Sciences, ECL normal for 3G10, 10E4 and 8G3 membranes and ECL prime for B-A38 membranes). Tubulin was used as a loading control. The membranes were digitalized using a GS-900 Calibrated Densitometer (BioRad). Two independent experiments were analyzed.

Glycosaminoglycan Enzymatic Digestion Assay

For HS enzymatic digestion, 125 µg of protein lysates of each sample were incubated with Heparinase III (EC 4.2.2.8; 5 mU/mL; Amsbio) in TBS 0.1% BSA, supplemented with 0,1 mM Calcium Acetate, overnight, at 37°C with constant mixing. CS enzymatic digestion was performed by incubating 50 µg of protein lysates of each sample with Chondroitinase ABC (EC 4.2.2.4; 10 mU/µL; Amsbio) in TBS 0.1% BSA, supplemented with 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$ overnight, at 37°C with constant mixing.

Real-time PCR analysis

For mRNA quantification/gene expression analysis, total RNA was extracted from MKN74, MKN45, MKN7, MKN1, AGS and NCI-N87 gastric cancer cell lines using TRIzol Reagent (Ambion). 5 µg of RNA from each cell condition was reverse transcribed into cDNA by incubation with random hexamers/primers, using the SuperScript™ IV Reverse

Transcriptase Kit (Invitrogen). Real-time PCR was performed using for each condition: 4 µL of cDNA diluted 1:10 in ultrapure water, 0.6 µL of each primer at 10 µM (Table 2), 10 µL of PowerUp™ SYBR™ Green Master Mix (Applied Biosystems) and ultrapure water up to a final volume of 20 µL. The real-time PCR run was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems). RQ values were determined for each gene and β-actin was used as a housekeeping gene to normalize relative gene expression ($RQ = 2^{-(\text{target Ct} - \beta\text{-actin Ct})}$). Two independent experiments with triplicates per condition were analyzed.

Table 2. *List of primers used in the real-time PCR analysis.*

Gene	Primer Fw	Primer Rv
SDC1	5'-ATGGCTCTGGGGATGACTCT-3'	5'-GTGGGAATAGCCGTCAGGAG-3'
SDC2	5'-AGGATGTAGAGAGTCCAGAGCT-3'	5'-TGTATCCTCTTCGGCTGGGT-3'
SDC3	5'-TGACATCCCTGAGAGGAGCA-3'	5'-GCTACCACCTCATTGGCTGT-3'
SDC4	5'-CCGGAGCCCTACCAGACGAT-3'	5'-AGGCACCAAGGGATGGACAA-3'
GPC1	5'-CATCGGGTGTGGAGAGTG-3'	5'-TGAGCGTGTCCCTGTTGTC-3'
GPC2	5'-CTGGGACACGACCTGGAC-3'	5'-GCCATCCAGTCATCTGCATAC-3'
GPC3	5'-CTGCTTCAGTCTGCAAGTATGG-3'	5'-GTGGAGTCAGGCTTGGGTAG -3'
GPC4	5'-AGTGTGGTCAGCGAACAGTG-3'	5'-CAAACATATCATTCAGGGATTTCTC-3'
GPC5	5'-GCCGCCCTGTAAGAACAC-3'	5'-TCATTCCATGCTTCTCTTTGC-3'
GPC6	5'-CCAGGCATAAGAAATTTGACG-3'	5'-CATGTACAGCATGCCATAGGTC-3'
EXTL2	5'-TGAACTGGAAACCAATGCAG-3'	5'-AGGAAATTGCTGCCAAACTG-3'
EXTL3	5'-CTCCGCCATCAGGAAATC-3'	5'-AGTTGGAGTTGTAGAGCCAGGA-3'
NDST1	5'-CTGCCCTCTACCTGTTCCTG-3'	5'-AACTGGATCTCCTCAAAGGTCTC-3'
NDST2	5'-CAAGAGCTGCGTACCAACC-3'	5'-GAGGGTCCGTGTGTAGTTCAG-3'
NDST3	5'-CCTTGCAGAAGAGATGTTTGG-3'	5'-GTAGCAGGATCAGTTCTTAGTTGTTG-3'
NDST4	5'-GACATTGGGCTCCATCTGAC-3'	5'-GCTGCTGTCCATCAATAATTAGC-3'
GLCE	5'-TGTGGAAGTCCGAGACAGAG-3'	5'-CTGGATTGGATAGAAATAGCCTTG-3'
HS2ST1	5'-TGGAGATGATTATAGACCAGGGTTAC-3'	5'-GCTATGGCCACAGAAGAACG-3'
HS6ST1	5'-GCAGGGAGTGGAGCTAACAG-3'	5'-AACAGTTCCAGTTCCCGAAA-3'
HS6ST2	5'-CGGTGCGATCTTCTCCAA-3'	5'-AGGACGATCACGGCAAATAG-3'
HS6ST3	5'-CAACCACAGCCACACCAG-3'	5'-CTTCTTCCATCACACATATGAAGAG-3'
HS3ST1	5'-CAGCCAGATGCCCTTCTC-3'	5'-AGACTCGCTCAGGCACTTTG-3'
HS3ST2	5'-GATTGGTACAGGAGCCTGATG-3'	5'-GGAGCCTCTTGAGTGACAAAG-3'
HS3ST3A	5'-GGCCGAGAGAACCTGAACTC-3'	5'-CGAGCGACAGTGACTTCCA-3'
HS3ST3B	5'-GCAGATCTTGCCTCGATGTC-3'	5'-GCGCACGAGTACAGGAACATA-3'
HS3ST4	5'-TAGAGCCGCACTTCTTCGAC-3'	5'-GGTTATTTGCCCATCCAAAG-3'
HS3ST5	5'-CATCCGGCAGTAGTCAAAGC-3'	5'-TTGTGATTTGCTGAGGGTAGG-3'
HS3ST6	5'-GCCCTGCTGGAGTTTCTG-3'	5'-GCGCTCGTAGCACCTGTC-3'
HSULF1	5'-CCAGCAGAAGCCAAAGAAAG-3'	5'-GAACGTGTCTGCCGAGTATG-3'
HSULF2	5'-GCCTGCAAGAGAAGGACAAG-3'	5'-AGCAGCTTGCGGAGTTTC-3'
CSGALNACT1	5'-GGAGACCCTGAACAATCCTG-3'	5'-GCCGTTTGAATTCGTGTTTG-3'
CSGALNACT2	5'-GCCATTGTTTATGCCAACCA-3'	5'-ATCCACCAATGGTCAGGAAA-3'
CHSY1	5'-GCCCAGAAATACCTGCAGAC-3'	5'-GCACTACTGGAATTGGTACAGATG-3'
CHPF	5'-GGTGCACTATAGCCATCTGGA-3'	5'-GGCACTTCGGAAATGAGG-3'
CHSY3	5'-GACTCAGTGTGTCTGGTCTTACG-3'	5'-TTGCTATTGTGAAGGTCTTGGA-3'
CHST11	5'-CGCTGCTGGAAGTGATGA-3'	5'-AGGATAAAGGATCCCAAGCAA-3'
CHST12	5'-GTAGCCGACAAATCCTTCCA-3'	5'-ACCGGTTTACCTCTGACTTGAC-3'
CHST13	5'-CCGGCATTTGGAAACAGA-3'	5'-TCCAGGTCATAGAGCTTCTGC-3'
CHST14	5'-CCACTGCCTAATGTCACCAA-3'	5'-ATGACAGGCAGAAGCACAGA-3'
CHST15	5'-GTGCCAGGAATAAAGTTCAACA-3'	5'-CACTGGATAAGTCCCGAGTGA-3'
CHST3	5'-TGCACAGCCTGAAGATGAGA-3'	5'-CAGCTTGTCTGAGACCCTTGA-3'
CHST7	5'-GATCCGGGTCAGTCACCA-3'	5'-GACAGATTGCCCCCACAG-3'
DSE	5'-GTCCAGAGGCACTTCAACATC-3'	5'-AGTCCGCAATAGCCACAGTC-3'
UST	5'-ACCATGGACCACCTCCTAGTAA-3'	5'-CACACTTGCCTACCCTGTTGTA-3'
β-actin	5′-AGAAAATCTGGCACCACACC-3′	5′- TAGCACAGCCTGGATAGCAA -3′

Clinical samples

For the tissue disaccharide structural analysis, human tissue samples were collected from gastric cancer surgical specimens. Tissue fragments were collected from both the cancer lesion and the area adjacent to the tumor, and were subjected to morphological quality control. The fragments were fast frozen in optimal cutting temperature compound (OCT) and kept at -80ºC until GAGs extraction.

Ethics Statement

Tissue samples were collected within the Tissue and Tumor Bank of Centro Hospitalar de S. João, which is approved by the Ethical Committee of Hospital de S. João, Porto (2005). All patients whose samples are stored in the Bank signed an informed consent.

GAGs extraction from tissues and purification

Tissue disaccharide structural analyses were conducted on GAGs extracted and purified from the tissue samples of 5 patients, both tumor tissue samples and distant mucosa tissue from the same individuals (m= <100mg per sample). The extraction and purification of GAGs was performed as described previously (408). Briefly, tissues were mechanically disrupted with a Potter grinder in a Urea denaturing buffer (50 mM TBS, EDTA 2mm, Urea 6M) and samples were centrifuged at 23000 g for 30 min at 4°C. The supernatant was recovered, and dialyzed in the dialysis buffer (Tris HCl 25 Mm, EDTA 5 mM, pH 7.8 – 8) for 24h at 4°C. Proteins were degraded by incubating samples with pronase (final concentration of 2mg/mL) for 24 h at 37° with stirring. Afterwards, samples were centrifuged at 23000g for 15min, supernatants were recovered and their protein content was precipitated by incubating with ice-cold trichloroacetic acid (final volume of 5% v/v) for 1 h, followed by additional centrifugation at 15000g for 20 min. Finally, sample supernatants were recovered and their lipidic fractions were removed through diethyl ether washing. For GAG purification, the extracted samples were applied to a DEAE-Sephacel column (Pharmacia Biotech) equilibrated in 20 mM phosphate pH 6.5. The column was then extensively washed with 20 mM phosphate pH 6.5, 0.3 M NaCl, and then GAGs were eluted with 20 mM phosphate pH 6.5, 1 M NaCl. Samples were desalted over a Pd-10 column (GE Healthcare), lyophilized, and stored at −20 °C prior to analysis.

GAG disaccharide structural analysis

For HS disaccharide analysis, samples were resuspended in 100 mM sodium acetate, 0.5 mM calcium acetate, pH 7.1, and then digested into disaccharides by incubation with a mix of heparinase I, II, and III (10 mU each) for 48 h at 37 °C. Compositional analysis was performed by reversed-phase ion pair HPLC, as described before (393). Briefly, samples

were applied to a C18 reversed phase column equilibrated in a H_2O/acetonitrile (8.5%) buffer supplemented with 1.2 mM of ion-pairing tetra-N-butylammonium hydrogen sulfate, then resolved using a multistep NaCl gradient calibrated with HS disaccharide standards. On-line post-column disaccharide derivatization was achieved by the addition of 2-cyanoacetamide (0.25%) in NaOH (0.5%), followed by fluorescence detection (excitation 346 nm, emission 410 nm). For CS disaccharide analysis, purified samples were digested into disaccharides by incubation with 500 mU of Chondroitinase ABC in 50 mM Tris–HCl pH 7.5, 50 mM NaCl, 2 mM $CaCl_2$, for 24 h at 37 °C. RPIP-HPLC was performed as described above for HS, with CS disaccharide standards used for calibration.

TCGA data analysis

mRNA gene expression data, quantified as FPKM (fragments per kilobase million) was retrieved from the Cancer Genome Atlas (TCGA) and comprised a total of 410 stomach adenocarcinoma samples and 38 normal gastric tissues. Clinical information from these samples was also downloaded from TCGA. Information on recurrence was downloaded from cBioPortal (https://www.cbioportal.org/).

Gene expression of different GAG-related genes was evaluated according to the available clinicopathological data, namely Lauren classification, following the information detailed in table 3. The “all cases” category included the data from *diffuse*, *intestinal* and *other* subtypes.

Table 3. *Gastric tumor tissue categories included in each subtype.*

Diffuse	Stomach, Adenocarcinoma, Diffuse Type
	Stomach Adenocarcinoma, Signet Ring Type
Intestinal	Stomach, Intestinal Adenocarcinoma, Not Otherwise Specified (NOS)
	Stomach, Intestinal Adenocarcinoma, Papillary Type
	Stomach, Intestinal Adenocarcinoma, Mucinous Type
	Stomach, Intestinal Adenocarcinoma, Tubular Type
Other	Stomach, Adenocarcinoma, Not Otherwise Specified (NOS)

Overall survival (OS) and disease-free survival (DFS) analysis were assessed using the KM test in R using the “survival” and “survminer” packages. Only individuals with survival information above zero were considered. Individuals were clustered into “high” and “low” expression groups based on the best separation cut-off method. This cut-off refers to the value, within the second quartile of expression of a given gene, that yields the lowest log-rank P-value. Only comparison with a log-rank P-value below 0.05 were considered.

Statistical Analysis

Statistical analysis was carried out using GraphPad Prism 6 and statistical significance was considered when P values were ≤0.05 (* means p≤0.05; ** means p≤0.01; *** means p≤0.001; **** means p≤0.0001). For real-time PCR, statistical significance was determined using unpaired Student's t-test with Welch's correction, with a 95% interval of confidence. In regards to the TCGA analysis, the statistical significance of the mRNA gene expression data was determined using ordinary one-way ANOVA with Tukey's multiple comparisons test.

Chapter 5

General Discussion

Future Perspectives and Concluding Remarks

1. General Discussion

Gastric cancer prevails as a concerning health issue, as the fifth leading cause of cancer related deaths worldwide (2). The lack of early detection screening methods and sensitive biomarkers, combined with the limited treatment options, that might not be effective for patients harboring different gastric cancer molecular profiles, configure significant challenges that must be overcome to improve patients' clinical outcomes (409). Remarkable scientific advances have been made over the last decades to better comprehend the molecular mechanisms underlying cell malignant transformation and cancer progression. Amongst many factors, protein glycosylation has been reported as a key event in cancer, where its disruption is recognized as a hallmark of the disease (51, 52, 55).

Protein glycosylation is a tightly regulated, non-template driven, process catalyzed by many different biosynthetic enzymes acting in a well-coordinated manner. The resulting glycoconjugates are prominent components of cell glycocalyx and the ECM, and mediate various cell signaling cascades, involved in proliferation, adhesion, motility, invasion, and angiogenesis. Alteration of the glycosylation pathways has been frequently reported and associated to the carcinogenic pathway in the vast majority of cancer pathologies, contributing to the formation of novel glycan epitopes or changes in the cellular glycans profile that are charged with pro-tumorigenic effects (52, 56). Therefore, glycans hold valuable potential as novel biomarkers and targets for clinical management (59).

Over the years, a few screening assays have been used in gastric cancer therapy based on the premise of glyco-targets, however limitations in sensitivity and specificity have restricted their applicability. Identifying the glycosylation alterations associating with the carcinogenic pathway and understanding its biological role in multiple malignant transformation events is critical to tackle the beforementioned clinical obstacles. A myriad of studies has highlighted the changes in HSPGs expression and cellular GAGosylation profiles as relevant players in tumorigenesis, particularly in gastric cancer (208, 270, 324, 326). Yet, the specific roles taken by these molecules within this specific context remain largely unexplored.

In this thesis we aimed to understand the general impact of HS-related changes in gastric cancer. We explored the role of multiple HS biosynthetic enzymes in the regulation of gastric cancer cell GAGosylation and glycoproteome profile, and further assessed the impact of HS in gastric cancer cell signaling and migration capabilities. Finally, we evaluated the clinical potential of HS and CS by evaluating their structural profile and GAG-associated gene signatures in gastric cancer, and correlating it with patient clinical features.

HS biosynthesis controls broad cellular GAGosylation and tumor cell aggressiveness

In **chapter 2**, we described the overall impact of HS expression in stomach cancer cell glycoprofile and behavior, by exploring the role of two glycosyltransferases that dictate HS biosynthesis, EXTL2 and EXTL3. We established glycoengineered models knocking-out either *EXTL2* or *EXTL3* in the MKN74 gastric cancer epithelial cell line*,* and evaluated its impact over GAGs biosynthesis*.* Consistent with previously published data, our work has demonstrated that **EXTL2** functions has a **negative regulator of HS biosynthesis and HSPGs expression,** as ***EXTL2* KO** contributed to increased **synthesis of HS** presenting **decreased 6-*O*-sulfation** motifs, **and increased *SDC4*** expression in tumor cells. On the other hand, we confirmed **EXTL3 key role** in **initiating HS biosynthesis** in **gastric cancer*.***

Ten years ago, a novel model was proposed for the regulation of HS biosynthesis mediated by EXTL2 and EXTL3, according to which, during the formation of the tetrasaccharide linker, the phosphorylation status of the initially added Xyl residue played a critical role in determining HS polymerization (Fig. 1). Following this model, Xyl phosphorylation by the kinase FAM20B is crucial for the completion of the tetrasaccharide linker. Nevertheless, Xyl dephosphorylation by the phosphatase XYLP, a step hypothesized to take place simultaneously with the addition of the first GlcA residue that finalizes the linker formation, is equally necessary for the polymerization of HS chains (144, 148). Therefore, similarly to what has been previously hypothesized (410), our data supports that EXTL2 enzymatic activity functions as a quality control mechanism, that inhibits polymerization of additional HS chains. It has been theorized that abnormally high *FAM20B* or low *XYLP* expression/activity can lead to the accumulation of phosphorylated tetrasaccharides (144). Similarly, if Xyl dephosphorylation by XYLP is a limiting step in this pathway, excessive increases in HS synthesis and impaired XYLP activity, could also promote the formation of phosphorylated tetrasaccharide linkers, to be targeted by EXTL2. By adding a GlcA residue to the phosphorylated linkers, EXTL2 would then catalyze the formation of phosphorylated pentasaccharides that cannot be further polymerized by the heterodimeric complex EXT1-EXT2, (148), nor dephosphorylated by XYLP and recycled into longer chains (144), inducing HS synthesis termination. This would result in the production of HSPGs modified with short incomplete glycan stubs (Fig. 1).

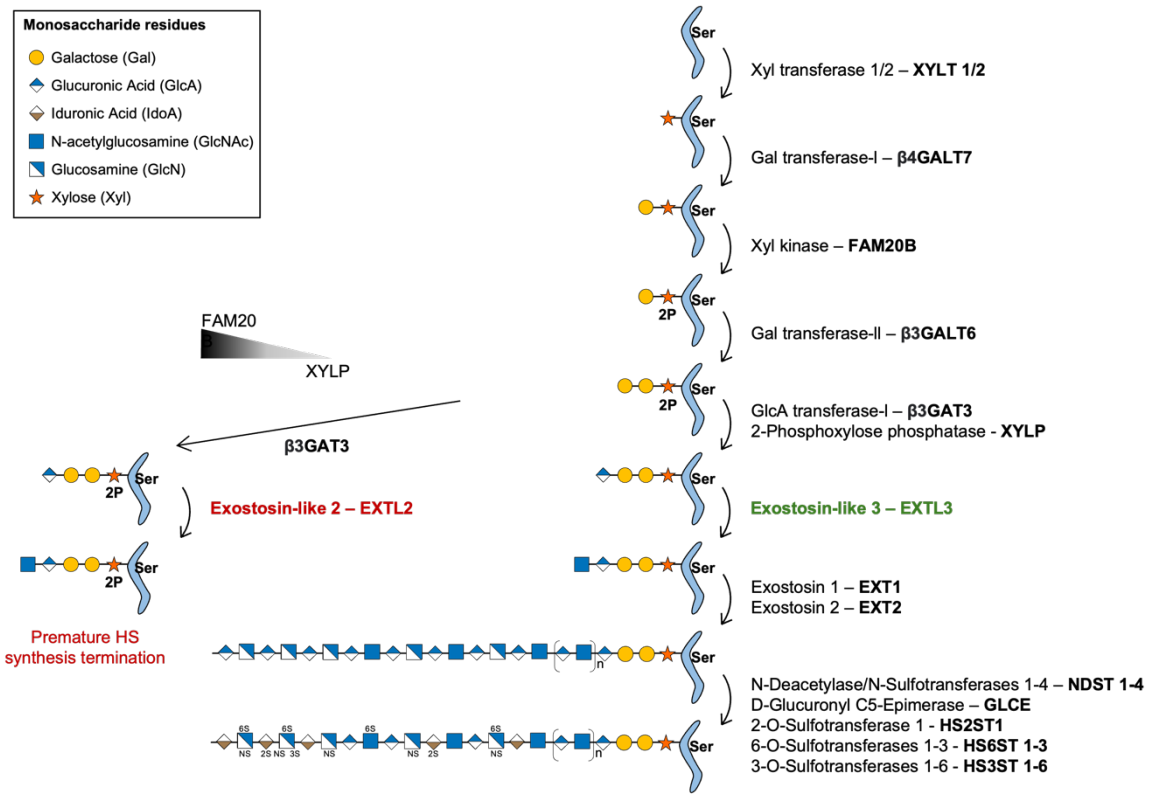


Figure 1. Schematic representation of HS biosynthesis depicting the proposed EXTL2- and EXTL3-mediated pathways.

Interestingly and in contradiction with the previous information, data from recent studies suggested that the linker phosphorylation might have no effect on the polymerization of EXTL3-primed linkers by EXT1-EXT2 (45, 356). This might allude to the cell-specificity of this process, as one of the studies was carried in Chinese Hamster Ovary (CHO) cells and the other one involved enzymological experiments carried in one-pot multienzyme (OPME) reaction conditions. Additionally, it highlights the need to further comprehend this specific mechanism in physiological conditions, and understand the regulatory role of EXTL2, distinct from EXTL3, in the absence of the linker phosphorylation.

Noteworthy, through detailed assessment of the impact of deregulated HS biosynthesis over membrane bound HSPGs, we demonstrated that besides promoting HS formation, ***EXTL2* KO increased *SDC4* transcript expression** and protein levels. In addition, WB analysis of glycoengineered *EXTL2/SDC4* dKO cell models revealed that **SDC4 is a main cell-surface HS-carrier** in MKN74 gastric cancer cells, further confirming the key role of this molecule in gastric tumors. Our data could support the hypothesis that EXTL2 regulates HS levels in MKN74 mainly by controlling the expression of SDC4. However, taking in consideration previous reports detailing EXTL2 enzymatic activity, it is more likely the existence of a EXTL2/SDC4 feedback axis, in which EXTL2 regulation of HS biosynthesis influences SDC4 expression. This would endow HS chains additional biological roles, namely the regulation of cell membrane HSPG expression and remodeling of cells

glycoproteome. In support of this hypothesis, a previous study has demonstrated that repression of HS biosynthesis in *extl3* mutant zebrafish embryos reduced significantly biglycan expression and showed a trend for increased aggrecan transcripts (364). Moreover, deregulated expression of the extracellular HS 6-*O*-endosulfatase *HSULF2*, and possibly altered HS sulfation motifs, were also reported to influence *GPC3* expression in liver cancer cells (322).

Additionally, our results support that *EXTL3* abrogation, and consequent blocking of HS biosynthesis, triggered **CS formation** and induced the expression of **CSPGs** and **HSPGs** abnormally modified with **CS chains,** including **SDC4.** There's a comprehensive body of research corroborating this HS:CS switch phenomenon that occurs frequently under impaired HS biosynthesis (356, 363), and targeting particularly SDC4 molecules (362, 364). Consistent with this information, our results support that abrogation of EXTL3 favors the accumulation of available tetrasaccharide linker substrates that, by lacking the first GlcNAc residue, are not polymerized by EXT1 and EXT2, and therefore are available to be primed for CS assembly by CS glycosyltransferases. It is important to state that, to our knowledge, this shift has been more frequently described under *in vitro* and *in vivo* conditions using models that have been genetically modified. The physiological and/or pathological triggers for PGs HS:CS ratio changes, and their associated biological implications remain largely unknown.

Knowing that HS and CS glycosyltransferases participate in the initial stages of either HS or CS polymerization by targeting the same tetrasaccharide linker structure, recently, a lot of interest has been garnered on the molecular mechanisms that determine the type of GAGs in the formation of HSPGs and CSPGs. Sammon *et al.* and Bourgeais *et al.* have shown that these enzymes have different substrate recognition and activity ratios. The CS glycosyltransferases tend to recognize the linker, rather than the modified polypeptide, and act in a more promiscuous manner, being able to target both HS- and CSPGs. On the other hand, HS glycosyltransferase EXTL3 interacts with the linker and the polypeptide backbone, targeting specific proteins determined to be HSPGs (45, 194). This decision is currently viewed as an all-encompassing process defined by several factors, including the spatiotemporal expression and co-localization of each GAG-specific glycosyltransferase, the levels of their respective UDP-sugar donors and the amount, co-localization and Golgi transit time of the membrane-bound core proteins. Therefore, it is easy to consider that even minor changes during this complex biosynthetic process might have impactful consequences in cells overall GAGosylation profile, rendering HS/CS biosynthesis as an interesting process to be assessed in disease.

The development and application of new technologies will significatively contribute to further enlighten the multi-factorial GAG biosynthesis process and the consequences of its

deregulation. As a confirmation of this, in a recent study, the combination of CRISPR-Cas9 knock-in technology and high-speed super-resolution microscopy allowed to acquire interesting data on the spatial resolution of different glycosylation enzymes in the Golgi (136). The work of Harada *et al.* allowed to distinguish the localization of early-stage enzymes of different glycosylation pathways (*N*- and *O*-glycosylation vs. GAG biosynthesis) within distinct areas of the same Golgi cisternae. Interestingly, it also demonstrated the highly dynamic movement of a GAG biosynthesis enzyme, XYLT2, through small tubules between Golgi cisternae, proposing Golgi fragmentation as a new molecular mechanism, subjacent to deregulated GAG biosynthesis, that could underly pathogenesis (136).

In **chapter 2,** we additionally described the biological impact of HS and GAGosylation changes in gastric cancer cell proliferation, viability, motility and invasion. The implications of HS and HSPGs over cell behavior have been previously reported in different types of tumor cells. Increased GPC3 levels in liver cancer associated with increased cell proliferative rates through the stimulation of the Hedgehog (Hh) signaling pathway (411). In prostate cancer cells, HSULF2 mediated changes of HS sulfation motifs increased cell viability and migration and expression of EMT markers, likely through the activation of Wnt signaling (412). In breast cancer, induced *HS2ST1* expression, and associated HS structural changes, reduced cell viability, migration and invasion by decreasing EGFR and E-cadherin levels, and inhibiting FGF2-induced MAPK activation (296). In our work, the phenotype resulting from either *EXTL2* KO or *EXTL3* KO did not elicit significant major changes in cell viability and proliferation. However, ***EXTL2* KO**, and concomitant HS and *SDC4* overexpression, induced an **aggressive behavior**, **increasing cell motility and invasion capabilities**. These data align with previous work reported by our team, showing that *SDC4* KO reduces significantly cell motility features and that its upregulation is correlated with gastric cancer patient poor prognosis (270). Similarly, Onyeisi *et al*. demonstrated that SDC4 silencing reduced breast tumor cells invasion and migration capabilities (269). Interestingly, an opposite trend has been reported in melanoma cells, where FGF2 stimulation reduced *SDC4* expression, which in turn was revealed to increase cell migration (413), re-enforcing once again the cell-specificity nature of HS and HSPG functions in pathology. Although the opposite effect would be expected in the absence of HS biosynthesis, **no significant changes in motility and invasion** were observed in ***EXTL3* KO cells**, suggesting that increased CS level and CS/HS functional overlap might have contributed to re-store cell behavior.

Finaly, in **chapter 2** we have described the influence of specific HS profiles in RTKs activation, aiming to uncover a signaling mechanism that could justify the aforementioned *EXTL2* KO cells' malignant behavior. Interestingly, our results **associated abrogation of *EXTL2* with overall decreased EphA 4 activation** (normalized for total receptor levels) in

gastric cancer cells. However, the major change observed for this receptor was mainly related to its protein levels, which were increased in *EXTL2* KO cells, rather than its phosphorylation. Interestingly EphA 4 upregulation has been previously reported as a bad prognostic factor in gastric cancer, although its biological role in the pathology remains to be elucidated (375-377).

Altered glycosylation in gastric cancer has been demonstrated to modulate RTKs activation, and consequently cell malignant transformation (77, 78). Taking this into consideration, and recognizing that most therapies currently applied in gastric cancer target RTKs, evaluating the role of HS in RTKs activation seemed to be a promising approach. Yet, despite the extensively described role of HS and HSPGs in the activation of cell surface RTKs in multiple types of cancer (94), our analyses revealed quite mild variations in the phosphorylation of most evaluated receptors, suggesting that different HS-mediated mechanisms might underly MKN74 cell motility and aggressive features. TGF-β mediated signaling is a potential pathway that might be involved in this process. TGF-β signaling has been implicated in tumor cell migration and invasion and in EMT, particularly in the context of gastric cancer (414). In addition, many reports have described an interesting association between TGF-β and SDC4. Ma *et al.* revealed a regulatory loop involving the RNA binding protein ZFP36L1, SDC4 and TGF-β in osteosarcoma, and showed that SDC4 promotes TGF-β signaling and TGF-β-induced tumor cell migration, and in turn its expression is also controlled by TGF-β signaling (262). In accordance, different reports have also described SDC4-mediated TGF-β signaling (415) and TGF-β-induced SDC4 expression (416). Similarly, our team has also described a positive correlation between SDC4 and TGF-β in gastric cancer cell derived EVs, which associated with EV-induced cell aggressive phenotype (270).

In **chapter 3**, we focused on the biological implications of HS fine structural features, namely the 6-*O*-sulfation motifs, regulated by the activity of the enzymes 6-*O*-sulfotransferase HS6ST1 and the 6-*O*-endosulfatase HSULF2. We established glycoengineered MKN74 gastric cancer cell models either knocking-out *HS6ST1* or inducing the OE of *HSULF2* and mutant *HSULF2ΔSG*. Our data revealed that the intracellular and extracellular regulation of HS 6-*O*-sulfation patterns influences overall HS biosynthesis mechanisms and the expression and, possibly membrane half-life, of cell surface HSPGs.

HS disaccharide structural analysis revealed that ***HS6ST1* KO** cells synthesized **HS chains lacking 6-*O*-sulfation motifs** and with increased *N*-sulfation and *N*- and 2-*O*-sulfation motifs. Since we did not observed variations in the expression of other enzymes participating in HS sulfation, including HS2ST1, we hypothesize that the observed changes

resulted from the accumulation of disaccharides that could not be 6-*O*-sulfated. Interestingly, WB and flow cytometry analysis demonstrated that repression of HS 6-*O*-sulfation, and the associated changes in HS sulfation profiles, **promoted the biosynthesis of HS**, and possibly **longer chains**, in this particular cell model. Since we did **not observe the same trend in *HSULF2* OE cells**, also associated to reduced HS 6-*O*-sulfation, our data support that the sulfation patterns impact HS length by fine-tuning its polymerization in the Golgi, and not by influencing the activity of post-synthesis modification enzymes that cleave HS chains, namely HPSE. In addition, the **expression profiles of HS glycosyltransferases** remained **unaltered in *HS6ST1* KO cells** supporting that the changes in HS 6-*O*-sulfation, *N*-sulfation and/or *N*- and 2-*O*-sulfation motifs may influence the substrate recognition and or enzymatic activity of HS polymerizing enzymes, EXT1 and EXT2. Another possible hypothesis is that HS6ST1 enzyme might directly interact or influence activity or localization of EXT1/2.

This intriguing relationship between HS sulfation and the fine-tuning of chains length has been previously described in the literature, as NDST2-mediated *N*-sulfation was shown to promote HS polymerization (391, 392). However, in a different study, no significant changes were determined in the length of HS chains isolated from embryonic fibroblasts from *Hs2st*$^{-/-}$ mutant mice, lacking 2-*O*-sulfation but harboring compensatory changes in 6-*O*- and *N*-sulfation (417). This suggests that specific sulfation motifs, rather than general sulfation and/or altered expression of specific sulfation enzymes, appear to be important for the regulation of HS polymerization. The cell type-specificity of HS regulation might also have a role in these differences.

These data are consistent with the currently accepted model for GAG biosynthesis, which proposes that GAG biosynthetic enzymes are in tight interaction with each other as part of a supramolecular complex, the “GAGosome” (134). There is a lot of evidence backing up this theory, and further detailing the influence of enzymes with different enzymatic activity over each other. McCormick *et al*. revealed that formation of a stable complex by EXT1 and EXT2 is important for its localization in the Golgi and polymerization of HS chains (150). Later, Presto J. and colleagues showed that NDST1 and EXT2 interact, and that *NDST1* expression and activity was positively influenced by the induced expression of *EXT2* in HEK293 cells, increasing HS *N*- and *O*-sulfation, whereas induced *EXT1* expression, alone or in combination with *EXT2*, resulted in the opposite effect over *NDST1* expression and HS *N*-sulfation (418). Overall, these data suggest the existence of dynamic complexes formed by these three enzymes whose varying expression and interactions ultimately impact HS structural motifs. The GlcA C5 epimerase and HS2ST1, two enzymes that catalyze subsequent steps in HS modification, were also revealed to physically interact with each other and to form an important complex that determines enzymes localization in

the Golgi and HS sulfation motifs (161, 419). Finally, it has also been reported possible interactions between the GlcA C5 epimerase and HS6STs, although the implications over HS features have not been determined (420).

Additionally, in **chapter 3,** we detailed the impact of deregulated HS biosynthesis, specifically in regards to HS sulfation, over the expression and protein levels of main cell surface HSPGs, namely SDCs and GPCs. qRT-PCR analysis showed that ***SDC4*** **expression was significantly increased in *HS6ST1* KO cells**, revealing a possible new mechanism for the regulation of this glycoconjugate in gastric cancer. Additionally, protein analysis indicated that a subset of these SDC4 molecules might be modified with **longer HS chains,** aligning with the previous results indicating increased HS content in *HS6ST1* KO cells. Finally, altered expression of either *HS6ST1* or *HSULF2* similarly increased SDC1 protein levels, without impact over its transcript levels, suggesting that **decreased HS 6-*O*-sulfation** might contribute to **increase the half-life of SDC1 in MKN74 cells**.

Finally, clinical transcriptomic data reported on **chapter 3** indicated that ***HSULF2*** **is overexpressed in gastric cancer** from both the intestinal and diffuse histological subtypes, and that it associates with **patient's poor disease-free survival**. These data supporting the pro-tumorigenic role of HSULF2 are in accordance with previous reports describing its upregulation in different types of cancer, including lung (318), pancreatic (319), breast (320), colorectal (321) and gastric cancer (335), and associating it with patients worst prognosis (322). Overall, it emphasizes the need to perform additional functional analyses to understand how HS 6-*O*-sulfation modulation (*HS6ST1 vs. HSULF2*) influences tumor cells aggressive behavior and unveil the associated oncogenic signaling cascades.

HS and SDC4: two sides of the same coin?

An increasing body of evidence has highlighted SDC4 as a fundamental glycoconjugate in gastric cell malignant transformation, particularly in the context of intestinal subtype gastric cancer. Our team has reported a significant increase of *SDC4* expression in the gastric mucosa of individuals infected by highly virulent strains of *H. pylori,* often associated with development of gastric pre-malignant conditions (327). More recently, we showed that *SDC4* is overexpressed on intestinal subtype gastric tumors and correlates with cancer cell aggressiveness features and patient worse prognosis (270).

Our data (**chapter 2 and 3**) supports that SDC4 is a main cell surface HS-carrier in intestinal subtype gastric cancer and, remarkably, it reveals that SDC4 expression might be tightly regulated by the structural features of its composing HS chains, prompting the question: are HS and SDC4 two sides of the same coin? Our data suggest that this might be the case, at least within the gastric context. Although our initial goal was to address specifically the impact of HS chains in this pathology, we could not disentangle the effect of

HS from SDC4 in cancer cell behavior, as SDC4 was a consistent indirect target of the induced changes in the expression of several HS biosynthesis enzymes. *EXTL2* KO and *HS6ST1* KO similarly remodulated MKN74 cells GAGosylation profile: both by increasing HS content and decreasing or abrogating 6-*O*-sulfation, respectively. Interestingly, these changes induced *SDC4* overexpression in both cell models.

Different authors have previously reported that PGs expression can be influenced by different growth factors. Jaakkola P. and colleagues have described cell-specific regulation of *SDC1* expression modulated by FGF2 in fibroblasts and by EGF in keratinocytes (421). More recently, as abovementioned, another report described *SDC4* reduction induced by FGF2 stimulation in melanoma cells (413). Additionally, incubation with the TGF-β1 was shown to substantially increase the expression of the CS/DS carrier biglycan (422).

Integrating these data with our results, we hypothesize that changes in HS structural features, including length or sulfation, can modulate HS-ligand binding affinity and change specific signaling pathways that ultimately modulate transcriptional changes that lead to overexpression of SDC4 modified with the structurally altered HS chains, further promoting this feedback mechanism (Fig. 2).

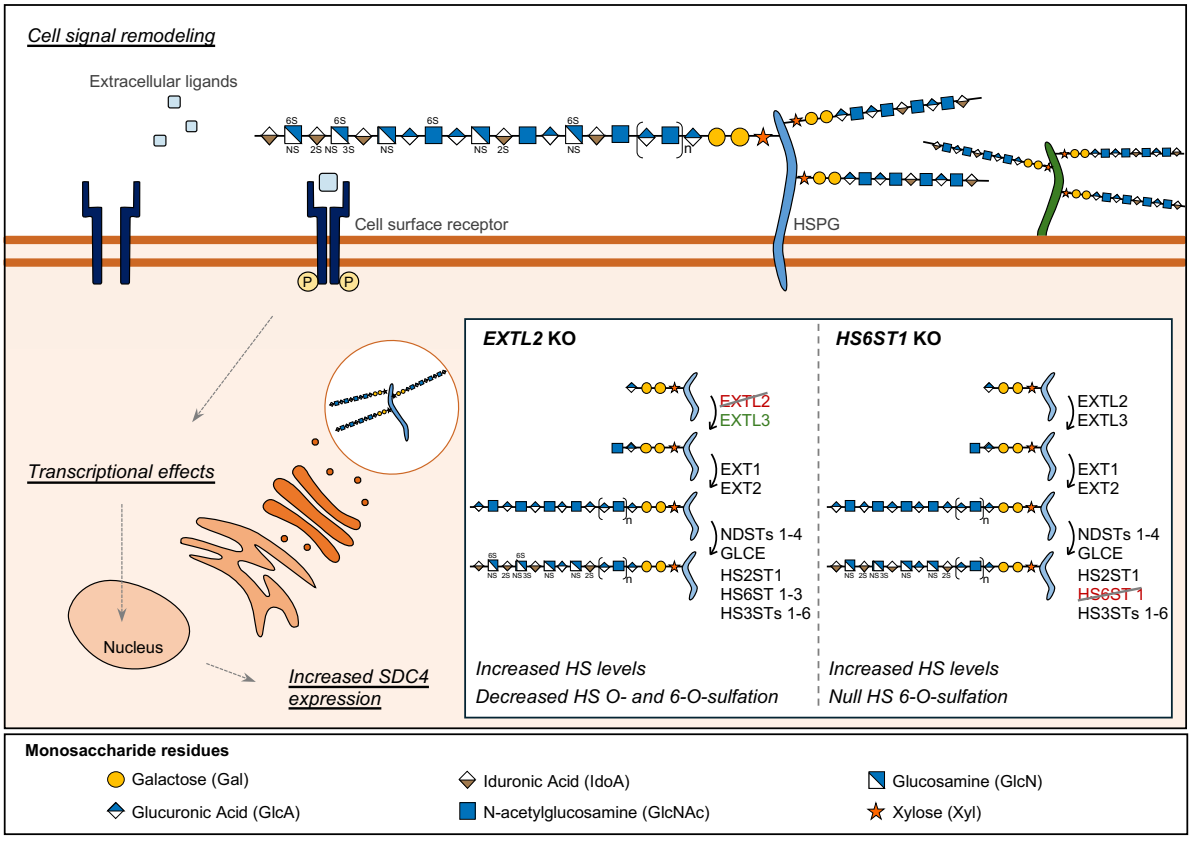


Figure 2. Schematic representation depicting the proposed feedback mechanism for the regulation of SDC4 expression by the induced HS structural features in the established MKN74 *EXTL2* KO and *HS6ST1* KO glycoengineered gastric cancer cell models.

Glycoengineered cell models as valuable cancer research glyco-tools

In our work, the use of glycoengineered cell models to study the role of HS structural features in gastric cancer (**chapter 2** and **3**) was crucial to characterize global GAG changes elicited by specific HS-related genes (both in HS and CS content) and to identify the above proposed regulatory mechanisms.

Over the years, many studies aimed to shed light on HS interactome and ligand binding specificities, and to further understand the implications and role of HS and HS biosynthesis on cell biology and in pathology. The use of heavily sulfated heparin molecules, that do not accurately represent HS structural features, and enzymatic assays, that disregard biological factors and the cell environment and variability, have been amongst the most frequently employed methods.

GAGs comprehend extremely large chains with highly variable cell-, tissue- and pathology-dependent structural features impacting their functions. The use of cell models and *in vitro* systems allow to study the biological functions of GAGs and GAG biosynthesis in different cell contexts, while still considering the structural diversity of GAGs, PGs core proteins and cellular distribution, and the many biological processes inherent to both physiology and pathology. By resorting to the CRISPR-Cas9 gene editing technology it is then possible to study the role of specific enzymes participating in GAG biosynthesis and its impact over cell GAGosylation profiles and signaling mechanisms (220, 356). Particularly in the context of cancer research, it further allows to study the regulation of cancer-associated GAGs and PGs and dissect its functional roles in different tumorigenic processes promoting disease development and progression.

There is still a considerable gap in knowledge regarding glycan structural cues that could help further understand GAG-protein interactions and the diverse roles of GAGs and PGs in different tissues and pathologies. These include GAG full sequences, the specific properties conferred by glycan chains' structural features, along with the specific functions elicited by these. Nonetheless, over the years we have witnessed the rise of relevant analytical glyco-tools that allow to study GAG chains as part of the PGs their attached to, and could help to address these challenges (79). Recently, Anggara *et al.* reported a sophisticated analytical approach to single-molecule analysis of glycoconjugates through direct, free-label, spatial imaging of individual glycoconjugates and their composing glycan chains. It was demonstrated that by employing low-temperature scanning tunneling microscopy, it was possible to reveal glycan structures at the monosaccharide level, including unveiling GAG chain sequences, along with the position/attachment site of these glycans in the glycoconjugates (423).

Identifying novel gastric cancer-associated GAG-signatures

Overall, **chapter 2** and **3** reported data showcasing that the mechanisms that dictate HS and HSPGs biosynthesis and structural motifs, and consequently define HS biological interactions and cellular functions, can be regulated by multiple factors. Changes in the expression of a single enzyme might translate into diverse and unexpected structural feature outcomes, highlighting the need to assess both structural and transcriptomic data when studying GAGosylation profiles and their roles in pathologies. This was our goal when planning the work described in **chapter 4.**

In **chapter 4,** we uncovered unique gastric cancer-associated GAGosylation profiles, by screening a set of gastric cancer cell lines, with different histopathological origin, and evaluating the expression of HS- and CS-related enzymes and HSPGs. qRT-PCR analysis revealed that the **six gastric cancer cell lines expressed the enzymatic machinery to carry all major steps of HS and CS/DS biosynthesis**. Regarding HS biosynthesis, *NDST1* and *NDST2*, the main NDST isoforms expressed in all vertebrate adult tissues, were detected in all cell lines. The epimerase *GLCE* and sulfotransferase *HS2ST1,* besides being expressed in all cells, exhibited a similar expression profile between the different cell lines, which is consistent with the known correlated role of these enzymes since HS2ST1 targets preferentially epimerized IdoA residues (159). *HS6ST1* was the main 6-O-sulfotransferase isoform expressed amongst all cell lines, consistent with the data described on chapter 3, reporting that abrogation of this isoform abolishes HS 6-*O*-sulfation. *HS3STs* expression, although low, varied between all cell lines, hinting at possible isoform-specific sulfation epitopes with distinctive biological implications in a cell type specific fashion (401, 404). *HSULF2,* commonly described as being pro-oncogenic, was detected in all cell lines, while *HSULF1* had a more restricted and cell-specific expression.

Regarding CS/DS biosynthesis, the expression of initiation and polymerization enzymes varied in all cell lines, with *CSGALNACT2*, *CHSY1* and *CHPF* being the predominantly expressed isoforms. The CS 4-*O*-sulfotransferases *CHST11-13* were detected in different cell lines, while CS 6-*O*-sulfotransferases, *CHST3*, *CHST7* and *CHST15,* were the only class of enzymes that was not expressed by all cell lines, as we did not detect transcript levels of any of these enzymes in MKN74 and AGS cells. The DS specific enzymes, *CHST14* and *DSE*, were expressed in all cells tested. Overall, this reduced expression of CS sulfotransferase enzymes in the gastric cancer cell lines is consistent with previous studies reporting a significant increase of non-sulfated CS/DS content in gastric cancer (325).

The HSPGs SDC1 (326), SDC4 (270) and GPC6 (328) are amongst the list of membrane bound PGs reported to be deregulated in gastric cancer. Therefore, we decided to focus on the expression of cell surface HSPGs. Transcriptomic analysis showed that

***SDC1*, *SDC4* and *GPC1* were the predominant HS-carriers expressed in all gastric cancer cell lines.** In addition, HS, HS stubs and CS (C4S and C6S stubs) cellular levels were also assessed. We observed significant variations of HS, HS stubs and C4S labeling amongst the cell lines, but could barely detect C6S stubs, which is consistent with the low or undetectable expression levels of CS 6-*O*-sulfotransferases in most cell lines. Additionally, by integrating the SDC4, HS stubs and C4S WB labeling profiles, we reported the existence of a **hybrid SDC molecule**, **possibly SDC4**, modified with HS and/or C4S **in NCI-N87 cells**. This shows that a hybrid SDC4 might be expressed in a tumor cell-dependent manner, without any artificial manipulation of the HS biosynthesis machinery. Nonetheless, this observation requires further validation, which could be achieved by performing GAG structural analyses of SDC4 molecules isolated from NCI-N87 cells.

In **chapter 4,** tumor-associated HS and CS/DS transcriptomic profiles and their correlation with patients' prognosis were analyzed by performing TCGA- STAD data mining and analysis. The gene expression data revealed an interesting trend suggesting **overall increased HS and CS formation, and decreased HS 6-*O*-sulfation in gastric cancer**, in both diffuse and intestinal subtypes. Interestingly a larger number of HS and CS/DS related genes influencing patients' outcome were detected in intestinal subtype gastric cancer. In intestinal gastric cancer, *SDC1* downregulation and *SDC4* and *GPC6* overexpression associated with patients' lower OS and DFS. In both intestinal and diffuse gastric cancer, downregulation of linker synthesis related genes mostly associated with lower DFS and OS. However, a similar trend in the association of decreased expression of HS polymerizing enzymes and patients' worse prognosis was only noticed in the intestinal subtype cases. No specific trend was reported when associating the expression of HS modification enzymes and patient's prognosis (great variation). Regarding CS biosynthesis, overexpression of different genes coding for CS sulfation and epimerization enzymes associated with lower OS, regardless of the disease subtype.

Noteworthy, we determined that several biosynthetic enzymes appear to be deregulated in the context of gastric cancer. This suggests that cancer patients GAG-profile likely results from the coordinated interaction of multiple components of the GAG enzymatic machinery that are deregulated during cancer progression, rather than a specific one, once again consistent with the concept of GAGosome. We further addressed this complexity in **chapter 4,** by optimization of a protocol for tissue GAG disaccharide structural analysis that allowed clinical samples GAG profiling. Over the years we have witnessed the rise of new analytical methodologies for effective and quick GAGs analysis. Structural disaccharide analyses, in particular, represent one of the most commonly used approaches for GAG detection and characterization, comprehending typically chromatography and mass-spectrometry approaches (79). In our work, we resorted to reversed-phase ion pairing

(RPIP)-HPLC (*via* fluorescence detection) based GAG disaccharide structural analysis to infer on the GAG content on tumor tissues and paired distant mucosa from gastric tumor patients. Our preliminary analyses, detailed in **chapter 4,** revealed that this methodology provides high resolution separation and sensitive detection of the different GAG disaccharides, allowing to evaluate the HS and CS/DS disaccharide content and sulfation composition from small tissue portions (<100 mg). We are now planning on increasing the patient's cohort, aiming to identify a specific gastric tumor-associated GAG structural signature, possibly overlooked in previous studies (208, 324, 325). Combining these structural analyses with tumor immunohistochemical data could provide additional insightful information on the GAG patterns on tumor microenvironment. Not only the levels but also the altered localization of GAG-modified PGs and the possibly altered GAGosylation profile of specific PGs could constitute important markers of disease onset and progression. The limitation of these analyses relies on the time-consuming tasks for sample preparation, including tissue GAG extraction and purification. This means, that although this is a reliable and sensitive technique to infer about specific disease associated-GAG patterns, further optimizations of this methodology, or different analytical tools would have to be employed to evaluate patient samples in the clinical context.

The application of different emergent methodologies like matrix-assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI), combining mass-spectrometry-based glycan structural analysis with histological data, would also provide means to detect, identify and assess the spatial localization of GAGs in patient tissues. Specifically in the context of gastric cancer, the use of a MALDI-MSI-based analytical approach in tissue micro-arrays has provided new insights on the differential distribution of GAG fragments in tissue sections, comparing tumor cells and stroma, and between different patients (424). Additionally, in the same study a specific GAG fragment (HexNAc–HexA–HexNAc) detected in the tumor stroma was reported to be an independent unfavorable prognostic factor for gastric cancer (424).

Another interesting approach holding high clinical potential for the development of diagnostic tools, pertains to the evaluation of EVs isolated from cancer patients' blood samples. EVs are in circulation in many body fluids, transporting different classes of macromolecules. Throughout recent years, several studies have described the presence and transport of GAGs and HSPGs in circulating EVs and highlighted the crucial role of these elements for EVs biogenesis, cargo, secretion, organotropism and uptake (111, 270, 425, 426). Deciphering specific GAG signatures of cancer-derived EVs would provide an important base for the development of non-invasive and sensitive screening serologic assays (427).

Recently, free GAG profiles of urine and plasma from healthy donors and cancer patients were analyzed, *via* high-throughput analytical screening, as potential non-invasive biomarkers. Changes in the profile of free GAGs were determined, even at early-stage cancer patients, across multiple cancer types. These profiles were associated with disease initiation and progression and patients' prognosis, further highlighting the potential of free GAGs screening for the development of a liquid biopsy platform for effective multi-cancer early detection (81).

Overall, there is an extensive body of research supporting not only the key role of GAGs in multiple cellular events underlying tumorigenesis, but also the clinical potential of GAG enzymatic machinery and GAG signatures. To assess the specific PG and GAG structural profiles that associate with disease is essential for the development of novel assays and tools to improve cancer early and effective diagnosis, to monitor disease progression and to undertake patient stratification for personalized therapy.

2. Future Perspectives and Concluding Remarks

The work detailed in this thesis led to the establishment and characterization of different glycoengineered gastric cancer cell models with specific GAGosylation profiles with functional implications in cell behavior. Additionally, we further identified a gastric cancer-associated GAG profile with impact in patient prognosis. In the future, we aim to address the following questions that have emerged from our work:

Which CSPGs are overexpressed in response to the deregulation of HS biosynthesis in MKN74 EXTL3 KO cell models?

Although we have been more focused on the effects of deregulated HS biosynthesis and SDCs expression in gastric cancer, our work showed interesting results hinting at possible implications of CS and CSPGs in the context of this pathology (**Chapter 2** - *Additional results*). We would like to further explore this route, starting by addressing the potential CSPGs overexpressed as a response mechanism to the lack of HS biosynthesis, by performing gene expression analysis.

What is the global influence of reduced HS 6-O-sulfation in tumor cell glycoproteome and PGs cellular distribution?

We were able to establish and characterize interesting gastric cancer cell models with altered sulfation profiles (**Chapter 2** and **3**). Our results showcased the general impact of altered HS biosynthesis mechanisms in cellular HS content, and unveiled a possible role for HS 6-*O*-sulfation in determining cells glycoproteome profile. Moving forward, to further study the length of HS chains in these models and validate these data, we plan to perform metabolic labeling and gel chromatography, and structural analysis of isolated SDC4 molecules. Additionally, we plan to further evaluate the impact of deregulated HS biosynthesis on HSPGs, by analyzing its content in cell glycocalyx and secretome. Flow cytometry, cell surface biotinylating and WB assays will be performed to achieve this objective.

What are the implications of reduced HS 6-O-sulfation in tumor cell behavior and oncogenic signaling cascades?

In addition to the previous point, we are planning to address the implications of *HS6ST1* KO and *HSULF2* OE induced HS and HSPG changes in cell behavior and cell signaling, by performing migration and invasion assays, and studying the activation of classical ligand-receptor axis known to be influenced by HS 6-*O*-sulfation motifs (e.g. FGF2 ligand stimulatory assay).

What are the molecular mechanisms underlying HSULF2 oncogenic role in gastric cancer?

TCGA data analysis revealed that *HSULF2* is upregulated in gastric cancer and that it correlates with patient's poor prognosis (**Chapter 3** and **4**). To further study the oncogenic implications of HS 6-*O*-sulfation fine tuning and *HSULF2* expression in gastric cancer, we have established CRISPR-Cas9 glycoengineered *HSULF2* KO cell models using the gastric cancer cell line MKN45, presenting increased HSULF2 mRNA and protein levels. These models are currently under biochemical characterization and will allow to study the opposite effects of *HSULF2* OE *vs. HSULF2* KO.

Is there a specific gastric cancer-associated GAG structural signature?

Finally, we plan to increase the patient cohort for tissue GAG disaccharide analysis in order to characterize the GAGosylation structural profiles of gastric cancer patients. In addition, by resorting to MALDI-imaging we aim to determine GAGs spatial resolution, comparing tumor and normal mucosa, and multiple areas of single tumor samples. To further uncover the possible changes in PGs GAGosylation profile, we also plan to perform Proximity Ligation Assay (PLA) analysis of CS and HS and different carriers in patient tissues along the gastric carcinogenesis pathway.

In summary, the work outlined in this thesis has shown that HS is involved in gastric cancer progression, endowing tumor cells with highly invasive and aggressive phenotypes. We showed that the cellular balance of the HS glycosyltransferases EXTL2 and EXTL3 dictates HS biosynthesis and the cellular HS:CS global ratio. Our work has further unveiled a possible regulatory network involving HS biosynthesis and structural motifs and the expression and glycosylation of SDC4, a main cell surface HSPG displaying pro-tumorigenic roles in gastric cancer.

References

1. Neelamegham, S., Aoki-Kinoshita, K., Bolton, E., Frank, M., Lisacek, F., Lütteke, T. *et al.* (2019) Updates to the Symbol Nomenclature for Glycans guidelines Glycobiology **29**, 620-624
2. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I. *et al.* (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries CA: A Cancer Journal for Clinicians **74**, 229-263
3. Iozzo, R. V., andSchaefer, L. (2015) Proteoglycan form and function: A comprehensive nomenclature of proteoglycans Matrix Biology **42**, 11-55
4. Merry, C. L. R., Lindahl, U., Couchman, J., andEsko, J. D. (2022) Proteoglycans and Sulfated Glycosaminoglycans In Essentials of Glycobiology, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., eds. Cold Spring Harbor Laboratory Press Copyright © 2022 The Consortium of Glycobiology Editors,, Cold Spring Harbor (NY) 217-232
5. Faria-Ramos, I., Poças, J., Marques, C., Santos-Antunes, J., Macedo, G., Reis, C. A. *et al.* (2021) Heparan Sulfate Glycosaminoglycans: (Un)Expected Allies in Cancer Clinical Management Biomolecules **11**,
6. de Visser, K. E., andJoyce, J. A. (2023) The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth Cancer Cell **41**, 374-403
7. Hanahan, D., andWeinberg, R. A. (2011) Hallmarks of cancer: the next generation Cell **144**, 646-674
8. Hanahan, D. (2022) Hallmarks of Cancer: New Dimensions Cancer Discovery **12**, 31-46
9. Colquhoun, A., Arnold, M., Ferlay, J., Goodman, K. J., Forman, D., andSoerjomataram, I. (2015) Global patterns of cardia and non-cardia gastric cancer incidence in 2012 Gut **64**, 1881-1888
10. Lordick, F., Carneiro, F., Cascinu, S., Fleitas, T., Haustermans, K., Piessen, G. *et al.* (2022) Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up Ann Oncol **33**, 1005-1020
11. Yusefi, A. R., Bagheri Lankarani, K., Bastani, P., Radinmanesh, M., andKavosi, Z. (2018) Risk Factors for Gastric Cancer: A Systematic Review Asian Pac J Cancer Prev **19**, 591-603
12. Karimi, P., Islami, F., Anandasabapathy, S., Freedman, N. D., andKamangar, F. (2014) Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention Cancer Epidemiol Biomarkers Prev **23**, 700-713
13. Wroblewski, L. E., Peek, R. M., Jr., andWilson, K. T. (2010) Helicobacter pylori and gastric cancer: factors that modulate disease risk Clin Microbiol Rev **23**, 713-739
14. Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M. *et al.* (2001) Helicobacter pylori infection and the development of gastric cancer N Engl J Med **345**, 784-789
15. David, S., andMeltzer, S. J. (2010) Stomach - Genetic and epigenetic alterations of preneoplastic and neoplastic lesions Cancer Biomark **9**, 493-507
16. Fukayama, M., Hayashi, Y., Iwasaki, Y., Chong, J., Ooba, T., Takizawa, T. *et al.* (1994) Epstein-Barr virus-associated gastric carcinoma and Epstein-Barr virus infection of the stomach Lab Invest **71**, 73-81
17. Carneiro, F., Oliveira, C., Suriano, G., andSeruca, R. (2008) Molecular pathology of familial gastric cancer, with an emphasis on hereditary diffuse gastric cancer J Clin Pathol **61**, 25-30
18. Lauren, P. (1965) The Two Histological Main Types Of Gastric Carcinoma: Diffuse And So-Called Intestinal-Type Carcinoma. An Attempt At A Histo-Clinical Classification Acta Pathol Microbiol Scand **64**, 31-49
19. Hu, B., El Hajj, N., Sittler, S., Lammert, N., Barnes, R., andMeloni-Ehrig, A. (2012) Gastric cancer: Classification, histology and application of molecular pathology J Gastrointest Oncol **3**, 251-261
20. Correa, P., andPiazuelo, M. B. (2012) The gastric precancerous cascade J Dig Dis **13**, 2-9
21. Carneiro, P., Fernandes, M. S., Figueiredo, J., Caldeira, J., Carvalho, J., Pinheiro, H. *et al.* (2012) E-cadherin dysfunction in gastric cancer - Cellular consequences, clinical applications and open questions FEBS Letters **586**, 2981-2989

22. Network, T. C. G. A. R. (2014) Comprehensive molecular characterization of gastric adenocarcinoma Nature **513**, 202-209
23. Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., andLordick, F. (2020) Gastric cancer Lancet **396**, 635-648
24. Duarte, H. O., Gomes, J., Machado, J. C., andReis, C. A. (2018) Gastric cancer: Basic aspects Helicobacter **23 Suppl 1**, e12523
25. Qi, C., Liu, C., Gong, J., Liu, D., Wang, X., Zhang, P. *et al.* (2024) Claudin18.2-specific CAR T cells in gastrointestinal cancers: phase 1 trial final results Nature Medicine
26. Wainberg, Z. A., Enzinger, P. C., Kang, Y. K., Qin, S., Yamaguchi, K., Kim, I. H. *et al.* (2022) Bemarituzumab in patients with FGFR2b-selected gastric or gastro-oesophageal junction adenocarcinoma (FIGHT): a randomised, double-blind, placebo-controlled, phase 2 study Lancet Oncol **23**, 1430-1440
27. Maron, S. B., Alpert, L., Kwak, H. A., Lomnicki, S., Chase, L., Xu, D. *et al.* (2018) Targeted Therapies for Targeted Populations: Anti-EGFR Treatment for EGFR-Amplified Gastroesophageal Adenocarcinoma Cancer Discov **8**, 696-713
28. Guan, W. L., He, Y., andXu, R. H. (2023) Gastric cancer treatment: recent progress and future perspectives J Hematol Oncol **16**, 57
29. Ferreira, J. A., Magalhães, A., Gomes, J., Peixoto, A., Gaiteiro, C., Fernandes, E. *et al.* (2017) Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics Cancer Letters **387**, 32-45
30. Shimada, H., Noie, T., Ohashi, M., Oba, K., andTakahashi, Y. (2014) Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the Task Force of the Japanese Gastric Cancer Association Gastric Cancer **17**, 26-33
31. Guadagni, F., Roselli, M., Amato, T., Cosimelli, M., Perri, P., Casale, V. *et al.* (1992) CA 72-4 measurement of tumor-associated glycoprotein 72 (TAG-72) as a serum marker in the management of gastric carcinoma Cancer Res **52**, 1222-1227
32. Nakagoe, T., Sawai, T., Tsuji, T., Jibiki, M., Nanashima, A., Yamaguchi, H. *et al.* (2001) Pre-operative serum levels of sialyl Tn antigen predict liver metastasis and poor prognosis in patients with gastric cancer Eur J Surg Oncol **27**, 731-739
33. Nakagoe, T., Sawai, T., Tsuji, T., Jibiki, M. A., Nanashima, A., Yamaguchi, H. *et al.* (2002) Predictive factors for preoperative serum levels of sialy Lewis(x), sialyl Lewis(a) and sialyl Tn antigens in gastric cancer patients Anticancer Res **22**, 451-458
34. Takahashi, Y., Takeuchi, T., Sakamoto, J., Touge, T., Mai, M., Ohkura, H. *et al.* (2003) The usefulness of CEA and/or CA19-9 in monitoring for recurrence in gastric cancer patients: a prospective clinical study Gastric Cancer **6**, 142-145
35. Duffy, M. J., Lamerz, R., Haglund, C., Nicolini, A., Kalousová, M., Holubec, L. *et al.* (2014) Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update Int J Cancer **134**, 2513-2522
36. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J., andClausen, H. (2020) Global view of human protein glycosylation pathways and functions Nature Reviews Molecular Cell Biology **21**, 729-749
37. Reily, C., Stewart, T. J., Renfrow, M. B., andNovak, J. (2019) Glycosylation in health and disease Nature Reviews Nephrology **15**, 346-366
38. Mereiter, S., Balmaña, M., Campos, D., Gomes, J., andReis, C. A. (2019) Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading? Cancer Cell **36**, 6-16
39. Flynn, R. A., Pedram, K., Malaker, S. A., Batista, P. J., Smith, B. A. H., Johnson, A. G. *et al.* (2021) Small RNAs are modified with N-glycans and displayed on the surface of living cells Cell **184**, 3109-3124.e3122
40. Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M. *et al.* (2022) In Essentials of Glycobiology, Cold Spring Harbor Laboratory Press Copyright © 2022 by the Consortium of Glycobiology Editors., Cold Spring Harbor (NY)
41. Magalhães, A., Duarte, H. O., andReis, C. A. (2021) The role of O-glycosylation in human disease Mol Aspects Med **79**, 100964

42. Bennett, E. P., Mandel, U., Clausen, H., Gerken, T. A., Fritz, T. A., andTabak, L. A. (2012) Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family Glycobiology **22**, 736-756
43. Hart, G. W. (2019) Nutrient regulation of signaling and transcription J Biol Chem **294**, 2211-2231
44. Toole, B. P. (2004) Hyaluronan: from extracellular glue to pericellular cue Nature Reviews Cancer **4**, 528-539
45. Sammon, D., Krueger, A., Busse-Wicher, M., Morgan, R. M., Haslam, S. M., Schumann, B. *et al.* (2023) Molecular mechanism of decision-making in glycosaminoglycan biosynthesis Nature Communications **14**, 6425
46. Ricard-Blum, S., Vivès, R. R., Schaefer, L., Götte, M., Merline, R., Passi, A. *et al.* (2024) A biological guide to glycosaminoglycans: current perspectives and pending questions The FEBS Journal **n/a**,
47. Basu, A., Patel, N. G., Nicholson, E. D., andWeiss, R. J. (2022) Spatiotemporal diversity and regulation of glycosaminoglycans in cell homeostasis and human disease Am J Physiol Cell Physiol **322**, C849-c864
48. Funderburgh, J. L. (2000) Keratan sulfate: structure, biosynthesis, and function Glycobiology **10**, 951-958
49. Caterson, B., andMelrose, J. (2018) Keratan sulfate, a complex glycosaminoglycan with unique functional capability Glycobiology **28**, 182-206
50. Varki, A. (2017) Biological roles of glycans Glycobiology **27**, 3-49
51. Christiansen, M. N., Chik, J., Lee, L., Anugraham, M., Abrahams, J. L., andPacker, N. H. (2014) Cell surface protein glycosylation in cancer Proteomics **14**, 525-546
52. Pinho, S. S., andReis, C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications Nature Reviews Cancer **15**, 540-555
53. Stowell, S. R., Ju, T., andCummings, R. D. (2015) Protein glycosylation in cancer Annu Rev Pathol **10**, 473-510
54. Peixoto, A., Relvas-Santos, M., Azevedo, R., Santos, L. L., andFerreira, J. A. (2019) Protein Glycosylation and Tumor Microenvironment Alterations Driving Cancer Hallmarks Front Oncol **9**, 380
55. Vajaria, B. N., andPatel, P. S. (2017) Glycosylation: a hallmark of cancer? Glycoconj J **34**, 147-156
56. Rodrigues, J. G., Balmaña, M., Macedo, J. A., Poças, J., Fernandes, Â., de-Freitas-Junior, J. C. M. *et al.* (2018) Glycosylation in cancer: Selected roles in tumour progression, immune modulation and metastasis Cell Immunol **333**, 46-57
57. Locker, G. Y., Hamilton, S., Harris, J., Jessup, J. M., Kemeny, N., Macdonald, J. S. *et al.* (2006) ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer J Clin Oncol **24**, 5313-5327
58. Diniz, F., Lamas, S., Osório, H., Aguiar, P., Freitas, D., Gärtner, F. *et al.* (2023) Nanoparticles targeting Sialyl-Tn for efficient tyrosine kinase inhibitor delivery in gastric cancer Acta Biomater **170**, 142-154
59. Mereiter, S., Balmaña, M., Gomes, J., Magalhães, A., andReis, C. A. (2016) Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer Front Oncol **6**, 55
60. Marcos, N. T., Bennett, E. P., Gomes, J., Magalhaes, A., Gomes, C., David, L. *et al.* (2011) ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues Front Biosci (Elite Ed) **3**, 1443-1455
61. Sewell, R., Bäckström, M., Dalziel, M., Gschmeissner, S., Karlsson, H., Noll, T. *et al.* (2006) The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer J Biol Chem **281**, 3586-3594
62. Gill, D. J., Clausen, H., andBard, F. (2011) Location, location, location: new insights into O-GalNAc protein glycosylation Trends Cell Biol **21**, 149-158

63. Ju, T., Lanneau, G. S., Gautam, T., Wang, Y., Xia, B., Stowell, S. R. *et al.* (2008) Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc Cancer Res **68**, 1636-1646
64. Werther, J. L., Rivera-MacMurray, S., Bruckner, H., Tatematsu, M., andItzkowitz, S. H. (1994) Mucin-associated sialosyl-Tn antigen expression in gastric cancer correlates with an adverse outcome Br J Cancer **69**, 613-616
65. Pinho, S., Marcos, N. T., Ferreira, B., Carvalho, A. S., Oliveira, M. J., Santos-Silva, F. *et al.* (2007) Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells Cancer Lett **249**, 157-170
66. Freitas, D., Campos, D., Gomes, J., Pinto, F., Macedo, J. A., Matos, R. *et al.* (2019) O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype EBioMedicine **40**, 349-362
67. Tian, H., Miyoshi, E., Kawaguchi, N., Shaker, M., Ito, Y., Taniguchi, N. *et al.* (2008) The implication of N-acetylglucosaminyltransferase V expression in gastric cancer Pathobiology **75**, 288-294
68. Taniguchi, N., andKorekane, H. (2011) Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics BMB Rep **44**, 772-781
69. Pinho, S. S., Reis, C. A., Paredes, J., Magalhães, A. M., Ferreira, A. C., Figueiredo, J. *et al.* (2009) The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin Hum Mol Genet **18**, 2599-2608
70. Alexander, K. L., Serrano, C. A., Chakraborty, A., Nearing, M., Council, L. N., Riquelme, A. *et al.* (2020) Modulation of glycosyltransferase ST6Gal-I in gastric cancer-derived organoids disrupts homeostatic epithelial cell turnover J Biol Chem **295**, 14153-14163
71. Rodrigues, J. G., Duarte, H. O., Gomes, C., Balmaña, M., Martins Á, M., Hensbergen, P. J. *et al.* (2021) Terminal α2,6-sialylation of epidermal growth factor receptor modulates antibody therapy response of colorectal cancer cells Cell Oncol (Dordr) **44**, 835-850
72. Duarte, H. O., Balmaña, M., Mereiter, S., Osório, H., Gomes, J., andReis, C. A. (2017) Gastric Cancer Cell Glycosylation as a Modulator of the ErbB2 Oncogenic Receptor International Journal of Molecular Sciences **18**, 2262
73. Duarte, H. O., Rodrigues, J. G., Gomes, C., Hensbergen, P. J., Ederveen, A. L. H., de Ru, A. H. *et al.* (2021) ST6Gal1 targets the ectodomain of ErbB2 in a site-specific manner and regulates gastric cancer cell sensitivity to trastuzumab Oncogene **40**, 3719-3733
74. Liu, N., Zhu, M., Linhai, Y., Song, Y., Gui, X., Tan, G. *et al.* (2018) Increasing HER2 α2,6 sialylation facilitates gastric cancer progression and resistance via the Akt and ERK pathways Oncol Rep **40**, 2997-3005
75. Baldus, S. E., Zirbes, T. K., Mönig, S. P., Engel, S., Monaca, E., Rafiqpoor, K. *et al.* (1998) Histopathological subtypes and prognosis of gastric cancer are correlated with the expression of mucin-associated sialylated antigens: Sialosyl-Lewis(a), Sialosyl-Lewis(x) and sialosyl-Tn Tumour Biol **19**, 445-453
76. Dall'Olio, F., Malagolini, N., Trinchera, M., andChiricolo, M. (2014) Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression Biochim Biophys Acta **1840**, 2752-2764
77. Mereiter, S., Magalhães, A., Adamczyk, B., Jin, C., Almeida, A., Drici, L. *et al.* (2016) Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer Biochim Biophys Acta **1860**, 1795-1808
78. Gomes, C., Osório, H., Pinto, M. T., Campos, D., Oliveira, M. J., andReis, C. A. (2013) Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells PLoS One **8**, e66737
79. Zappe, A., Miller, R. L., Struwe, W. B., andPagel, K. (2022) State-of-the-art glycosaminoglycan characterization Mass Spectrom Rev **41**, 1040-1071
80. Gatto, F., Volpi, N., Nilsson, H., Nookaew, I., Maruzzo, M., Roma, A. *et al.* (2016) Glycosaminoglycan Profiling in Patients' Plasma and Urine Predicts the Occurrence of Metastatic Clear Cell Renal Cell Carcinoma Cell Rep **15**, 1822-1836

81. Bratulic, S., Limeta, A., Dabestani, S., Birgisson, H., Enblad, G., Stålberg, K. *et al.* (2022) Noninvasive detection of any-stage cancer using free glycosaminoglycans Proceedings of the National Academy of Sciences **119**, e2115328119
82. Rodriguez, E., Lindijer, D. V., van Vliet, S. J., Garcia Vallejo, J. J., andvan Kooyk, Y. (2024) The transcriptional landscape of glycosylation-related genes in cancer iScience **27**, 109037
83. Esko, J. D., Kimata, K., andLindahl, U. (2009) Proteoglycans and sulfated glycosaminoglycans Essentials of Glycobiology 2nd edition
84. Couchman, J. R., Multhaupt, H., andSanderson, R. D. (2016) Recent Insights into Cell Surface Heparan Sulphate Proteoglycans and Cancer F1000Res **5**,
85. Theocharis, A. D., Skandalis, S. S., Gialeli, C., andKaramanos, N. K. (2016) Extracellular matrix structure Adv Drug Deliv Rev **97**, 4-27
86. Sarrazin, S., Lamanna, W. C., andEsko, J. D. (2011) Heparan sulfate proteoglycans Cold Spring Harb Perspect Biol **3**,
87. Karamanos, N. K., Piperigkou, Z., Theocharis, A. D., Watanabe, H., Franchi, M., Baud, S. *et al.* (2018) Proteoglycan Chemical Diversity Drives Multifunctional Cell Regulation and Therapeutics Chem Rev **118**, 9152-9232
88. Matsuo, I., andKimura-Yoshida, C. (2014) Extracellular distribution of diffusible growth factors controlled by heparan sulfate proteoglycans during mammalian embryogenesis Philos Trans R Soc Lond B Biol Sci **369**,
89. Hull, E. E., Montgomery, M. R., andLeyva, K. J. (2017) Epigenetic Regulation of the Biosynthesis & Enzymatic Modification of Heparan Sulfate Proteoglycans: Implications for Tumorigenesis and Cancer Biomarkers Int J Mol Sci **18**,
90. Annaval, T., Wild, R., Crétinon, Y., Sadir, R., Vivès, R. R., andLortat-Jacob, H. (2020) Heparan Sulfate Proteoglycans Biosynthesis and Post Synthesis Mechanisms Combine Few Enzymes and Few Core Proteins to Generate Extensive Structural and Functional Diversity Molecules **25**,
91. Pessentheiner, A. R., Ducasa, G. M., andGordts, P. (2020) Proteoglycans in Obesity-Associated Metabolic Dysfunction and Meta-Inflammation Front Immunol **11**, 769
92. De Pasquale, V., Quiccione, M. S., Tafuri, S., Avallone, L., andPavone, L. M. (2021) Heparan Sulfate Proteoglycans in Viral Infection and Treatment: A Special Focus on SARS-CoV-2 Int J Mol Sci **22**,
93. O'Callaghan, P., Zhang, X., andLi, J. P. (2018) Heparan Sulfate Proteoglycans as Relays of Neuroinflammation J Histochem Cytochem **66**, 305-319
94. Hassan, N., Greve, B., Espinoza-Sánchez, N. A., andGötte, M. (2021) Cell-surface heparan sulfate proteoglycans as multifunctional integrators of signaling in cancer Cell Signal **77**, 109822
95. Noborn, F., Nilsson, J., andLarson, G. (2022) Site-specific glycosylation of proteoglycans: A revisited frontier in proteoglycan research Matrix Biol **111**, 289-306
96. Kolset, S. O., andPejler, G. (2011) Serglycin: A Structural and Functional Chameleon with Wide Impact on Immune Cells The Journal of Immunology **187**, 4927-4933
97. Korpetinou, A., Skandalis, S. S., Labropoulou, V. T., Smirlaki, G., Noulas, A., Karamanos, N. K. *et al.* (2014) Serglycin: at the crossroad of inflammation and malignancy Front Oncol **3**, 327
98. Forsberg, E., Pejler, G., Ringvall, M., Lunderius, C., Tomasini-Johansson, B., Kusche-Gullberg, M. *et al.* (1999) Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme Nature **400**, 773-776
99. Meen, A. J., Øynebråten, I., Reine, T. M., Duelli, A., Svennevig, K., Pejler, G. *et al.* (2011) Serglycin is a major proteoglycan in polarized human endothelial cells and is implicated in the secretion of the chemokine GROalpha/CXCL1 J Biol Chem **286**, 2636-2647
100. Zhang, L., Yang, M., Yang, D., Cavey, G., Davidson, P., andGibson, G. (2010) Molecular interactions of MMP-13 C-terminal domain with chondrocyte proteins Connect Tissue Res **51**, 230-239

101. Gondelaud, F., andRicard-Blum, S. (2019) Structures and interactions of syndecans Febs j **286**, 2994-3007
102. Chen, J., Wang, F., He, C., andLuo, S.-Z. (2021) Multiple dimerizing motifs at different locations modulate the dimerization of the syndecan transmembrane domains Journal of Molecular Graphics and Modelling **106**, 107938
103. Zimmermann, P., andDavid, G. (1999) The syndecans, tuners of transmembrane signaling Faseb j **13 Suppl**, S91-s100
104. Couchman, J. R., Gopal, S., Lim, H. C., Nørgaard, S., andMulthaupt, H. A. (2015) Fell-Muir Lecture: Syndecans: from peripheral coreceptors to mainstream regulators of cell behaviour Int J Exp Pathol **96**, 1-10
105. Afratis, N. A., Nikitovic, D., Multhaupt, H. A. B., Theocharis, A. D., Couchman, J. R., andKaramanos, N. K. (2017) Syndecans – key regulators of cell signaling and biological functions The FEBS Journal **284**, 27-41
106. Shworak, N. W., Shirakawa, M., Mulligan, R. C., andRosenberg, R. D. (1994) Characterization of ryudocan glycosaminoglycan acceptor sites Journal of Biological Chemistry **269**, 21204-21214
107. Deepa, S. S., Yamada, S., Zako, M., Goldberger, O., andSugahara, K. (2004) Chondroitin sulfate chains on syndecan-1 and syndecan-4 from normal murine mammary gland epithelial cells are structurally and functionally distinct and cooperate with heparan sulfate chains to bind growth factors. A novel function to control binding of midkine, pleiotrophin, and basic fibroblast growth factor J Biol Chem **279**, 37368-37376
108. Steinfeld, R., Van Den Berghe, H., andDavid, G. (1996) Stimulation of fibroblast growth factor receptor-1 occupancy and signaling by cell surface-associated syndecans and glypican J Cell Biol **133**, 405-416
109. Robinson, C. J., Mulloy, B., Gallagher, J. T., andStringer, S. E. (2006) VEGF165-binding sites within heparan sulfate encompass two highly sulfated domains and can be liberated by K5 lyase J Biol Chem **281**, 1731-1740
110. Alexopoulou, A. N., Multhaupt, H. A. B., andCouchman, J. R. (2007) Syndecans in wound healing, inflammation and vascular biology The International Journal of Biochemistry & Cell Biology **39**, 505-528
111. Baietti, M. F., Zhang, Z., Mortier, E., Melchior, A., Degeest, G., Geeraerts, A. *et al.* (2012) Syndecan–syntenin–ALIX regulates the biogenesis of exosomes Nature Cell Biology **14**, 677-685
112. Christianson, H. C., andBelting, M. (2014) Heparan sulfate proteoglycan as a cell-surface endocytosis receptor Matrix Biol **35**, 51-55
113. Wang, H., Jin, H., Beauvais, D. M., andRapraeger, A. C. (2014) Cytoplasmic domain interactions of syndecan-1 and syndecan-4 with α6β4 integrin mediate human epidermal growth factor receptor (HER1 and HER2)-dependent motility and survival J Biol Chem **289**, 30318-30332
114. Teng, Y. H.-F., Aquino, R. S., andPark, P. W. (2012) Molecular functions of syndecan-1 in disease Matrix Biology **31**, 3-16
115. Yoneda, A., andCouchman, J. R. (2003) Regulation of cytoskeletal organization by syndecan transmembrane proteoglycans Matrix Biology **22**, 25-33
116. Filmus, J. (2023) Glypicans, 35 years later Proteoglycan Research **1**, e5
117. Kaur, S. P., andCummings, B. S. (2019) Role of glypicans in regulation of the tumor microenvironment and cancer progression Biochemical Pharmacology **168**, 108-118
118. Pomin, V. H., andMulloy, B. (2018) Glycosaminoglycans and Proteoglycans Pharmaceuticals **11**, 27
119. Filmus, J., Capurro, M., andRast, J. (2008) Glypicans Genome Biol **9**, 224
120. Filmus, J., andCapurro, M. (2014) The role of glypicans in Hedgehog signaling Matrix Biol **35**, 248-252
121. Iozzo, R. V., Zoeller, J. J., andNyström, A. (2009) Basement membrane proteoglycans: modulators Par Excellence of cancer growth and angiogenesis Mol Cells **27**, 503-513

122. Whitelock, J. M., Melrose, J., andIozzo, R. V. (2008) Diverse cell signaling events modulated by perlecan Biochemistry **47**, 11174-11183
123. Gubbiotti, M. A., Neill, T., andIozzo, R. V. (2017) A current view of perlecan in physiology and pathology: A mosaic of functions Matrix Biology **57-58**, 285-298
124. Hayes, A. J., Farrugia, B. L., Biose, I. J., Bix, G. J., andMelrose, J. (2022) Perlecan, A Multi-Functional, Cell-Instructive, Matrix-Stabilizing Proteoglycan With Roles in Tissue Development Has Relevance to Connective Tissue Repair and Regeneration Frontiers in Cell and Developmental Biology **10**,
125. Bezakova, G., andRuegg, M. A. (2003) New insights into the roles of agrin Nature Reviews Molecular Cell Biology **4**, 295-309
126. Dong, S., Cole, G. J., andHalfter, W. (2003) Expression of collagen XVIII and localization of its glycosaminoglycan attachment sites J Biol Chem **278**, 1700-1707
127. Iozzo, R. V. (1998) Matrix proteoglycans: from molecular design to cellular function Annu Rev Biochem **67**, 609-652
128. Iozzo, R. V., Goldoni, S., Berendsen, A. D., andYoung, M. F. (2011) Small Leucine-Rich Proteoglycans In The Extracellular Matrix: an Overview, Mecham RP, ed. Springer Berlin Heidelberg, Berlin, Heidelberg 197-231
129. Delehedde, M., Devenyns, L., Maurage, C. A., andVivès, R. R. (2013) Endocan in cancers: a lesson from a circulating dermatan sulfate proteoglycan Int J Cell Biol **2013**, 705027
130. Béchard, D., Scherpereel, A., Hammad, H., Gentina, T., Tsicopoulos, A., Aumercier, M. *et al.* (2001) Human endothelial-cell specific molecule-1 binds directly to the integrin CD11a/CD18 (LFA-1) and blocks binding to intercellular adhesion molecule-1 J Immunol **167**, 3099-3106
131. Shin, J. W., Huggenberger, R., andDetmar, M. (2008) Transcriptional profiling of VEGF-A and VEGF-C target genes in lymphatic endothelium reveals endothelial-specific molecule-1 as a novel mediator of lymphangiogenesis Blood **112**, 2318-2326
132. Lindahl, U. (2014) A personal voyage through the proteoglycan field Matrix Biology **35**, 3-7
133. Zhang, X., Wang, F., andSheng, J. (2016) "Coding" and "Decoding": hypothesis for the regulatory mechanism involved in heparan sulfate biosynthesis Carbohydr Res **428**, 1-7
134. Esko, J. D., andSelleck, S. B. (2002) Order out of chaos: assembly of ligand binding sites in heparan sulfate Annu Rev Biochem **71**, 435-471
135. Li, J. P., andKusche-Gullberg, M. (2016) Heparan Sulfate: Biosynthesis, Structure, and Function Int Rev Cell Mol Biol **325**, 215-273
136. Harada, A., Kunii, M., Kurokawa, K., Sumi, T., Kanda, S., Zhang, Y. *et al.* (2024) Dynamic movement of the Golgi unit and its glycosylation enzyme zones Nature Communications **15**, 4514
137. Meneghetti, M. C. Z., Deboni, P., Palomino, C. M. V., Braga, L. P., Cavalheiro, R. P., Viana, G. M. *et al.* (2021) ER-Golgi dynamics of HS-modifying enzymes via vesicular trafficking is a critical prerequisite for the delineation of HS biosynthesis Carbohydrate Polymers **255**, 117477
138. Götting, C., Kuhn, J., Zahn, R., Brinkmann, T., andKleesiek, K. (2000) Molecular cloning and expression of human UDP-d-Xylose:proteoglycan core protein beta-d-xylosyltransferase and its first isoform XT-II J Mol Biol **304**, 517-528
139. Almeida, R., Levery, S. B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E. P. *et al.* (1999) Cloning and expression of a proteoglycan UDP-galactose:beta-xylose beta1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family J Biol Chem **274**, 26165-26171
140. Koike, T., Izumikawa, T., Tamura, J., andKitagawa, H. (2009) FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan-protein linkage region Biochem J **421**, 157-162
141. Wen, J., Xiao, J., Rahdar, M., Choudhury, B. P., Cui, J., Taylor, G. S. *et al.* (2014) Xylose phosphorylation functions as a molecular switch to regulate proteoglycan biosynthesis Proc Natl Acad Sci U S A **111**, 15723-15728

142. Bai, X., Zhou, D., Brown, J. R., Crawford, B. E., Hennet, T., andEsko, J. D. (2001) Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6) J Biol Chem **276**, 48189-48195
143. Kitagawa, H., Taoka, M., Tone, Y., andSugahara, K. (2001) Human glycosaminoglycan glucuronyltransferase I gene and a related processed pseudogene: genomic structure, chromosomal mapping and characterization Biochem J **358**, 539-546
144. Koike, T., Izumikawa, T., Sato, B., andKitagawa, H. (2014) Identification of phosphatase that dephosphorylates xylose in the glycosaminoglycan-protein linkage region of proteoglycans J Biol Chem **289**, 6695-6708
145. Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J. C., Mulliert, G., Netter, P. *et al.* (2005) Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human beta1,4-galactosyltransferase 7 (GalT-I) and beta1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans J Biol Chem **280**, 1417-1425
146. Tone, Y., Pedersen, L. C., Yamamoto, T., Izumikawa, T., Kitagawa, H., Nishihara, J. *et al.* (2008) 2-o-phosphorylation of xylose and 6-o-sulfation of galactose in the protein linkage region of glycosaminoglycans influence the glucuronyltransferase-I activity involved in the linkage region synthesis J Biol Chem **283**, 16801-16807
147. Prydz , K. (2015) Determinants of Glycosaminoglycan (GAG) Structure Biomolecules **5**, 2003-2022
148. Nadanaka, S., Zhou, S., Kagiyama, S., Shoji, N., Sugahara, K., Sugihara, K. *et al.* (2013) EXTL2, a member of the EXT family of tumor suppressors, controls glycosaminoglycan biosynthesis in a xylose kinase-dependent manner J Biol Chem **288**, 9321-9333
149. Kim, B. T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U. *et al.* (2001) Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4-N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/ heparin biosynthesis Proceedings of the National Academy of Sciences of the United States of America **98**, 7176-7181
150. McCormick, C., Duncan, G., Goutsos, K. T., andTufaro, F. (2000) The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate Proc Natl Acad Sci U S A **97**, 668-673
151. Kitagawa, H., Shimakawa, H., andSugahara, K. (1999) The tumor suppressor EXT-like gene EXTL2 encodes an alpha1, 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region. The key enzyme for the chain initiation of heparan sulfate J Biol Chem **274**, 13933-13937
152. Busse, M., andKusche-Gullberg, M. (2003) In vitro polymerization of heparan sulfate backbone by the EXT proteins J Biol Chem **278**, 41333-41337
153. Senay, C., Lind, T., Muguruma, K., Tone, Y., Kitagawa, H., Sugahara, K. *et al.* (2000) The EXT1/EXT2 tumor suppressors: catalytic activities and role in heparan sulfate biosynthesis EMBO Rep **1**, 282-286
154. Aikawa, J., Grobe, K., Tsujimoto, M., andEsko, J. D. (2001) Multiple isozymes of heparan sulfate/heparin GlcNAc N-deacetylase/GlcN N-sulfotransferase. Structure and activity of the fourth member, NDST4 J Biol Chem **276**, 5876-5882
155. Li, Y. J., Yin, F. X., Zhang, X. K., Yu, J., Zheng, S., Song, X. L. *et al.* (2018) Characterization of heparan sulfate N-deacetylase/N-sulfotransferase isoform 4 using synthetic oligosaccharide substrates Biochim Biophys Acta Gen Subj **1862**, 547-556
156. Debarnot, C., Monneau, Y. R., Roig-Zamboni, V., Delauzun, V., Le Narvor, C., Richard, E. *et al.* (2019) Substrate binding mode and catalytic mechanism of human heparan sulfate d-glucuronyl C5 epimerase Proc Natl Acad Sci U S A **116**, 6760-6765
157. Hagner-Mcwhirter, A., Lindahl, U., andLi, J. (2000) Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5 Biochem J **347 Pt 1**, 69-75

158. Sheng, J., Xu, Y., Dulaney, S. B., Huang, X., andLiu, J. (2012) Uncovering biphasic catalytic mode of C5-epimerase in heparan sulfate biosynthesis J Biol Chem **287**, 20996-21002
159. Rong, J., Habuchi, H., Kimata, K., Lindahl, U., andKusche-Gullberg, M. (2001) Substrate specificity of the heparan sulfate hexuronic acid 2-O-sulfotransferase Biochemistry **40**, 5548-5555
160. Jacobsson, I., Lindahl, U., Jensen, J. W., Rodén, L., Prihar, H., andFeingold, D. S. (1984) Biosynthesis of heparin. Substrate specificity of heparosan N-sulfate D-glucuronosyl 5-epimerase J Biol Chem **259**, 1056-1063
161. Préchoux, A., Halimi, C., Simorre, J. P., Lortat-Jacob, H., andLaguri, C. (2015) C5-epimerase and 2-O-sulfotransferase associate in vitro to generate contiguous epimerized and 2-O-sulfated heparan sulfate domains ACS Chem Biol **10**, 1064-1071
162. Liu, C., Sheng, J., Krahn, J. M., Perera, L., Xu, Y., Hsieh, P. H. *et al.* (2014) Molecular mechanism of substrate specificity for heparan sulfate 2-O-sulfotransferase J Biol Chem **289**, 13407-13418
163. El Masri, R., Seffouh, A., Lortat-Jacob, H., andVivès, R. R. (2017) The “in and out” of glucosamine 6-O-sulfation: the 6th sense of heparan sulfate Glycoconjugate Journal **34**, 285-298
164. Smeds, E., Habuchi, H., Do, A. T., Hjertson, E., Grundberg, H., Kimata, K. *et al.* (2003) Substrate specificities of mouse heparan sulphate glucosaminyl 6-O-sulphotransferases Biochem J **372**, 371-380
165. Jemth, P., Smeds, E., Do, A. T., Habuchi, H., Kimata, K., Lindahl, U. *et al.* (2003) Oligosaccharide library-based assessment of heparan sulfate 6-O-sulfotransferase substrate specificity J Biol Chem **278**, 24371-24376
166. Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K. *et al.* (2000) The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine J Biol Chem **275**, 2859-2868
167. Jacobsson, I., andLindahl, U. (1980) Biosynthesis of heparin. Concerted action of late polymer-modification reactions J Biol Chem **255**, 5094-5100
168. Thieker, D. F., Xu, Y., Chapla, D., Nora, C., Qiu, H., Felix, T. *et al.* (2018) Downstream Products are Potent Inhibitors of the Heparan Sulfate 2-O-Sulfotransferase Sci Rep **8**, 11832
169. Kusche, M., Bäckström, G., Riesenfeld, J., Petitou, M., Choay, J., andLindahl, U. (1988) Biosynthesis of heparin. O-sulfation of the antithrombin-binding region J Biol Chem **263**, 15474-15484
170. Zhang, L., Lawrence, R., Schwartz, J. J., Bai, X., Wei, G., Esko, J. D. *et al.* (2001) The effect of precursor structures on the action of glucosaminyl 3-O-sulfotransferase-1 and the biosynthesis of anticoagulant heparan sulfate J Biol Chem **276**, 28806-28813
171. Liu, J., Shworak, N. W., Sinaÿ, P., Schwartz, J. J., Zhang, L., Fritze, L. M. *et al.* (1999) Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities J Biol Chem **274**, 5185-5192
172. Liu, J., Shriver, Z., Blaiklock, P., Yoshida, K., Sasisekharan, R., andRosenberg, R. D. (1999) Heparan sulfate D-glucosaminyl 3-O-sulfotransferase-3A sulfates N-unsubstituted glucosamine residues J Biol Chem **274**, 38155-38162
173. Lawrence, R., Yabe, T., Hajmohammadi, S., Rhodes, J., McNeely, M., Liu, J. *et al.* (2007) The principal neuronal gD-type 3-O-sulfotransferases and their products in central and peripheral nervous system tissues Matrix Biol **26**, 442-455
174. Xu, D., Tiwari, V., Xia, G., Clement, C., Shukla, D., andLiu, J. (2005) Characterization of heparan sulphate 3-O-sulphotransferase isoform 6 and its role in assisting the entry of herpes simplex virus type 1 Biochem J **385**, 451-459
175. Chen, J., Duncan, M. B., Carrick, K., Pope, R. M., andLiu, J. (2003) Biosynthesis of 3-O-sulfated heparan sulfate: unique substrate specificity of heparan sulfate 3-O-sulfotransferase isoform 5 Glycobiology **13**, 785-794

176. Vlodavsky, I., Ilan, N., andSanderson, R. D. (2020) Forty Years of Basic and Translational Heparanase Research Adv Exp Med Biol **1221**, 3-59
177. Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R. *et al.* (1999) Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis Nat Med **5**, 793-802
178. Wu, L., Viola, C. M., Brzozowski, A. M., andDavies, G. J. (2015) Structural characterization of human heparanase reveals insights into substrate recognition Nature Structural & Molecular Biology **22**, 1016-1022
179. Masola, V., Bellin, G., Gambaro, G., andOnisto, M. (2018) Heparanase: A Multitasking Protein Involved in Extracellular Matrix (ECM) Remodeling and Intracellular Events Cells **7**,
180. Vlodavsky, I., Singh, P., Boyango, I., Gutter-Kapon, L., Elkin, M., Sanderson, R. D. *et al.* (2016) Heparanase: From basic research to therapeutic applications in cancer and inflammation Drug Resist Updat **29**, 54-75
181. McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R. *et al.* (2000) Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member Biochem Biophys Res Commun **276**, 1170-1177
182. Levy-Adam, F., Feld, S., Cohen-Kaplan, V., Shteingauz, A., Gross, M., Arvatz, G. *et al.* (2010) Heparanase 2 interacts with heparan sulfate with high affinity and inhibits heparanase activity J Biol Chem **285**, 28010-28019
183. Kiyan, Y., Tkachuk, S., Kurselis, K., Shushakova, N., Stahl, K., Dawodu, D. *et al.* (2019) Heparanase-2 protects from LPS-mediated endothelial injury by inhibiting TLR4 signalling Sci Rep **9**, 13591
184. Kayal, Y., Barash, U., Naroditsky, I., Ilan, N., andVlodavsky, I. (2023) Heparanase 2 (Hpa2)-a new player essential for pancreatic acinar cell differentiation Cell Death & Disease **14**, 465
185. Pempe, E. H., Burch, T. C., Law, C. J., andLiu, J. (2012) Substrate specificity of 6-O-endosulfatase (Sulf-2) and its implications in synthesizing anticoagulant heparan sulfate Glycobiology **22**, 1353-1362
186. Seffouh, A., Milz, F., Przybylski, C., Laguri, C., Oosterhof, A., Bourcier, S. *et al.* (2013) HSulf sulfatases catalyze processive and oriented 6-O-desulfation of heparan sulfate that differentially regulates fibroblast growth factor activity Faseb j **27**, 2431-2439
187. El Masri, R., Crétinon, Y., Gout, E., andVivès, R. R. (2020) HS and Inflammation: A Potential Playground for the Sulfs? Front Immunol **11**, 570
188. Morimoto-Tomita, M., Uchimura, K., Werb, Z., Hemmerich, S., andRosen, S. D. (2002) Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans J Biol Chem **277**, 49175-49185
189. Holst, C. R., Bou-Reslan, H., Gore, B. B., Wong, K., Grant, D., Chalasani, S. *et al.* (2007) Secreted sulfatases Sulf1 and Sulf2 have overlapping yet essential roles in mouse neonatal survival PLoS One **2**, e575
190. Lamanna, W. C., Frese, M. A., Balleininger, M., andDierks, T. (2008) Sulf loss influences N-, 2-O-, and 6-O-sulfation of multiple heparan sulfate proteoglycans and modulates fibroblast growth factor signaling J Biol Chem **283**, 27724-27735
191. Zhang, L., David, G., andEsko, J. D. (1995) Repetitive Ser-Gly sequences enhance heparan sulfate assembly in proteoglycans J Biol Chem **270**, 27127-27135
192. Zhang, L., andEsko, J. D. (1994) Amino acid determinants that drive heparan sulfate assembly in a proteoglycan J Biol Chem **269**, 19295-19299
193. Chen, R. L., andLander, A. D. (2001) Mechanisms Underlying Preferential Assembly of Heparan Sulfate on Glypican-1* Journal of Biological Chemistry **276**, 7507-7517
194. Bourgeais, M., Fouladkar, F., Weber, M., Boeri-Erba, E., andWild, R. (2024) Chemo-enzymatic synthesis of tetrasaccharide linker peptides to study the divergent step in glycosaminoglycan biosynthesis Glycobiology **34**,
195. Sato, T., Gotoh, M., Kiyohara, K., Akashima, T., Iwasaki, H., Kameyama, A. *et al.* (2003) Differential roles of two N-acetylgalactosaminyltransferases, CSGalNAcT-1, and a novel

enzyme, CSGalNAcT-2. Initiation and elongation in synthesis of chondroitin sulfate J Biol Chem **278**, 3063-3071
196. Izumikawa, T., Koike, T., Shiozawa, S., Sugahara, K., Tamura, J., andKitagawa, H. (2008) Identification of chondroitin sulfate glucuronyltransferase as chondroitin synthase-3 involved in chondroitin polymerization: chondroitin polymerization is achieved by multiple enzyme complexes consisting of chondroitin synthase family members J Biol Chem **283**, 11396-11406
197. Izumikawa, T., Uyama, T., Okuura, Y., Sugahara, K., andKitagawa, H. (2007) Involvement of chondroitin sulfate synthase-3 (chondroitin synthase-2) in chondroitin polymerization through its interaction with chondroitin synthase-1 or chondroitin-polymerizing factor Biochem J **403**, 545-552
198. Kitagawa, H., Izumikawa, T., Uyama, T., andSugahara, K. (2003) Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization J Biol Chem **278**, 23666-23671
199. Mikami, T., andKitagawa, H. (2013) Biosynthesis and function of chondroitin sulfate Biochim Biophys Acta **1830**, 4719-4733
200. Mizumoto, S., andYamada, S. (2021) An Overview of in vivo Functions of Chondroitin Sulfate and Dermatan Sulfate Revealed by Their Deficient Mice Front Cell Dev Biol **9**, 764781
201. Khan, S., Fung, K. W., Rodriguez, E., Patel, R., Gor, J., Mulloy, B. *et al.* (2013) The solution structure of heparan sulfate differs from that of heparin: implications for function J Biol Chem **288**, 27737-27751
202. Lamanna, W. C., Kalus, I., Padva, M., Baldwin, R. J., Merry, C. L., andDierks, T. (2007) The heparanome--the enigma of encoding and decoding heparan sulfate sulfation J Biotechnol **129**, 290-307
203. Warda, M., Toida, T., Zhang, F., Sun, P., Munoz, E., Xie, J. *et al.* (2006) Isolation and characterization of heparan sulfate from various murine tissues Glycoconj J **23**, 555-563
204. Garner, O. B., Bush, K. T., Nigam, K. B., Yamaguchi, Y., Xu, D., Esko, J. D. *et al.* (2011) Stage-dependent regulation of mammary ductal branching by heparan sulfate and HGF-cMet signaling Dev Biol **355**, 394-403
205. Shah, M. M., Sakurai, H., Gallegos, T. F., Sweeney, D. E., Bush, K. T., Esko, J. D. *et al.* (2011) Growth factor-dependent branching of the ureteric bud is modulated by selective 6-O sulfation of heparan sulfate Dev Biol **356**, 19-27
206. Nigam, S. K., andBush, K. T. (2014) Growth factor-heparan sulfate "switches" regulating stages of branching morphogenesis Pediatr Nephrol **29**, 727-735
207. Jayson, G. C., Lyon, M., Paraskeva, C., Turnbull, J. E., Deakin, J. A., andGallagher, J. T. (1998) Heparan sulfate undergoes specific structural changes during the progression from human colon adenoma to carcinoma in vitro J Biol Chem **273**, 51-57
208. Theocharis, A. D., Vynios, D. H., Papageorgakopoulou, N., Skandalis, S. S., andTheocharis, D. A. (2003) Altered content composition and structure of glycosaminoglycans and proteoglycans in gastric carcinoma The International Journal of Biochemistry & Cell Biology **35**, 376-390
209. Fernández-Vega, I., García, O., Crespo, A., Castañón, S., Menéndez, P., Astudillo, A. *et al.* (2013) Specific genes involved in synthesis and editing of heparan sulfate proteoglycans show altered expression patterns in breast cancer BMC Cancer **13**, 24
210. Billings, P. C., andPacifici, M. (2015) Interactions of signaling proteins, growth factors and other proteins with heparan sulfate: mechanisms and mysteries Connect Tissue Res **56**, 272-280
211. Jakobsson, L., Kreuger, J., Holmborn, K., Lundin, L., Eriksson, I., Kjellén, L. *et al.* (2006) Heparan sulfate in trans potentiates VEGFR-mediated angiogenesis Dev Cell **10**, 625-634
212. Gopal, S., Søgaard, P., Multhaupt, H. A., Pataki, C., Okina, E., Xian, X. *et al.* (2015) Transmembrane proteoglycans control stretch-activated channels to set cytosolic calcium levels J Cell Biol **210**, 1199-1211

213. Casu, B., Naggi, A., andTorri, G. (2010) Heparin-derived heparan sulfate mimics to modulate heparan sulfate-protein interaction in inflammation and cancer Matrix Biol **29**, 442-452
214. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., andOrnitz, D. M. (1991) Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor Cell **64**, 841-848
215. Rapraeger, A. C., Krufka, A., andOlwin, B. B. (1991) Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation Science **252**, 1705-1708
216. Xie, Y., Su, N., Yang, J., Tan, Q., Huang, S., Jin, M. *et al.* (2020) FGF/FGFR signaling in health and disease Signal Transduct Target Ther **5**, 181
217. Saksela, O., Moscatelli, D., Sommer, A., andRifkin, D. B. (1988) Endothelial cell-derived heparan sulfate binds basic fibroblast growth factor and protects it from proteolytic degradation J Cell Biol **107**, 743-751
218. Maccarana, M., Casu, B., andLindahl, U. (1993) Minimal sequence in heparin/heparan sulfate required for binding of basic fibroblast growth factor J Biol Chem **268**, 23898-23905
219. Sugaya, N., Habuchi, H., Nagai, N., Ashikari-Hada, S., andKimata, K. (2008) 6-O-sulfation of heparan sulfate differentially regulates various fibroblast growth factor-dependent signalings in culture J Biol Chem **283**, 10366-10376
220. Qiu, H., Shi, S., Yue, J., Xin, M., Nairn, A. V., Lin, L. *et al.* (2018) A mutant-cell library for systematic analysis of heparan sulfate structure-function relationships Nat Methods **15**, 889-899
221. Pye, D. A., Vives, R. R., Turnbull, J. E., Hyde, P., andGallagher, J. T. (1998) Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity J Biol Chem **273**, 22936-22942
222. Kreuger, J., Salmivirta, M., Sturiale, L., Giménez-Gallego, G., andLindahl, U. (2001) Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2 J Biol Chem **276**, 30744-30752
223. Takei, Y., Ozawa, Y., Sato, M., Watanabe, A., andTabata, T. (2004) Three Drosophila EXT genes shape morphogen gradients through synthesis of heparan sulfate proteoglycans Development **131**, 73-82
224. Han, C., Belenkaya, T. Y., Khodoun, M., Tauchi, M., Lin, X., andLin, X. (2004) Distinct and collaborative roles of Drosophila EXT family proteins in morphogen signalling and gradient formation Development **131**, 1563-1575
225. Kleinschmit, A., Koyama, T., Dejima, K., Hayashi, Y., Kamimura, K., andNakato, H. (2010) Drosophila heparan sulfate 6-O endosulfatase regulates Wingless morphogen gradient formation Dev Biol **345**, 204-214
226. Gao, W., Xu, Y., Liu, J., andHo, M. (2016) Epitope mapping by a Wnt-blocking antibody: evidence of the Wnt binding domain in heparan sulfate Sci Rep **6**, 26245
227. Smith, R. A. A., Murali, S., Rai, B., Lu, X., Lim, Z. X. H., Lee, J. J. L. *et al.* (2018) Minimum structural requirements for BMP-2-binding of heparin oligosaccharides Biomaterials **184**, 41-55
228. Thacker, B. E., Xu, D., Lawrence, R., andEsko, J. D. (2014) Heparan sulfate 3-O-sulfation: a rare modification in search of a function Matrix Biol **35**, 60-72
229. Mulloy, B., Hogwood, J., Gray, E., Lever, R., andPage, C. P. (2016) Pharmacology of Heparin and Related Drugs Pharmacol Rev **68**, 76-141
230. Lindahl, U., Bäckström, G., Thunberg, L., andLeder, I. G. (1980) Evidence for a 3-O-sulfated D-glucosamine residue in the antithrombin-binding sequence of heparin Proc Natl Acad Sci U S A **77**, 6551-6555
231. Lindahl, U., Thunberg, L., Bäckström, G., Riesenfeld, J., Nordling, K., andBjörk, I. (1984) Extension and structural variability of the antithrombin-binding sequence in heparin J Biol Chem **259**, 12368-12376
232. Atha, D. H., Lormeau, J. C., Petitou, M., Rosenberg, R. D., andChoay, J. (1987) Contribution of 3-O- and 6-O-sulfated glucosamine residues in the heparin-induced conformational change in antithrombin III Biochemistry **26**, 6454-6461

233. Ye, S., Luo, Y., Lu, W., Jones, R. B., Linhardt, R. J., Capila, I. *et al.* (2001) Structural basis for interaction of FGF-1, FGF-2, and FGF-7 with different heparan sulfate motifs Biochemistry **40**, 14429-14439
234. Vanpouille, C., Deligny, A., Delehedde, M., Denys, A., Melchior, A., Liénard, X. *et al.* (2007) The heparin/heparan sulfate sequence that interacts with cyclophilin B contains a 3-O-sulfated N-unsubstituted glucosamine residue J Biol Chem **282**, 24416-24429
235. Thacker, B. E., Seamen, E., Lawrence, R., Parker, M. W., Xu, Y., Liu, J. *et al.* (2016) Expanding the 3-O-Sulfate Proteome--Enhanced Binding of Neuropilin-1 to 3-O-Sulfated Heparan Sulfate Modulates Its Activity ACS Chem Biol **11**, 971-980
236. Lindahl, U., andKjellén, L. (2013) Pathophysiology of heparan sulphate: many diseases, few drugs Journal of Internal Medicine **273**, 555-571
237. Piperigkou, Z., Mohr, B., Karamanos, N., andGötte, M. (2016) Shed proteoglycans in tumor stroma Cell Tissue Res **365**, 643-655
238. De Pasquale, V., andPavone, L. M. (2020) Heparan Sulfate Proteoglycan Signaling in Tumor Microenvironment International Journal of Molecular Sciences **21**, 6588
239. Leivonen, M., Lundin, J., Nordling, S., von Boguslawski, K., andHaglund, C. (2004) Prognostic value of syndecan-1 expression in breast cancer Oncology **67**, 11-18
240. Barbareschi, M., Maisonneuve, P., Aldovini, D., Cangi, M. G., Pecciarini, L., Angelo Mauri, F. *et al.* (2003) High syndecan-1 expression in breast carcinoma is related to an aggressive phenotype and to poorer prognosis Cancer **98**, 474-483
241. Hilgers, K., Ibrahim, S. A., Kiesel, L., Greve, B., Espinoza-Sánchez, N. A., andGötte, M. (2022) Differential Impact of Membrane-Bound and Soluble Forms of the Prognostic Marker Syndecan-1 on the Invasiveness, Migration, Apoptosis, and Proliferation of Cervical Cancer Cells Front Oncol **12**, 803899
242. Conejo, J. R., Kleeff, J., Koliopanos, A., Matsuda, K., Zhu, Z. W., Goecke, H. *et al.* (2000) Syndecan-1 expression is up-regulated in pancreatic but not in other gastrointestinal cancers Int J Cancer **88**, 12-20
243. Juuti, A., Nordling, S., Lundin, J., Louhimo, J., andHaglund, C. (2005) Syndecan-1 expression--a novel prognostic marker in pancreatic cancer Oncology **68**, 97-106
244. Nikolova, V., Koo, C. Y., Ibrahim, S. A., Wang, Z., Spillmann, D., Dreier, R. *et al.* (2009) Differential roles for membrane-bound and soluble syndecan-1 (CD138) in breast cancer progression Carcinogenesis **30**, 397-407
245. Nassar, E., Hassan, N., El-Ghonaimy, E. A., Hassan, H., Abdullah, M. S., Rottke, T. V. *et al.* (2021) Syndecan-1 Promotes Angiogenesis in Triple-Negative Breast Cancer through the Prognostically Relevant Tissue Factor Pathway and Additional Angiogenic Routes Cancers (Basel) **13**,
246. Ibrahim, S. A., Gadalla, R., El-Ghonaimy, E. A., Samir, O., Mohamed, H. T., Hassan, H. *et al.* (2017) Syndecan-1 is a novel molecular marker for triple negative inflammatory breast cancer and modulates the cancer stem cell phenotype via the IL-6/STAT3, Notch and EGFR signaling pathways Mol Cancer **16**, 57
247. Matsumoto, A., Ono, M., Fujimoto, Y., Gallo, R. L., Bernfield, M., andKohgo, Y. (1997) Reduced expression of syndecan-1 in human hepatocellular carcinoma with high metastatic potential Int J Cancer **74**, 482-491
248. Inki, P., Joensuu, H., Grénman, R., Klemi, P., andJalkanen, M. (1994) Association between syndecan-1 expression and clinical outcome in squamous cell carcinoma of the head and neck Br J Cancer **70**, 319-323
249. Kim, S. Y., Choi, E. J., Yun, J. A., Jung, E. S., Oh, S. T., Kim, J. G. *et al.* (2015) Syndecan-1 expression is associated with tumor size and EGFR expression in colorectal carcinoma: a clinicopathological study of 230 cases Int J Med Sci **12**, 92-99
250. Hua, R., Yu, J., Yan, X., Ni, Q., Zhi, X., Li, X. *et al.* (2020) Syndecan-2 in colorectal cancer plays oncogenic role via epithelial-mesenchymal transition and MAPK pathway Biomed Pharmacother **121**, 109630

251. Tsoyi, K., Osorio, J. C., Chu, S. G., Fernandez, I. E., De Frias, S. P., Sholl, L. *et al.* (2019) Lung Adenocarcinoma Syndecan-2 Potentiates Cell Invasiveness Am J Respir Cell Mol Biol **60**, 659-666
252. De Oliveira, T., Abiatari, I., Raulefs, S., Sauliunaite, D., Erkan, M., Kong, B. *et al.* (2012) Syndecan-2 promotes perineural invasion and cooperates with K-ras to induce an invasive pancreatic cancer cell phenotype Mol Cancer **11**, 19
253. Lim, H. C., Multhaupt, H. A. B., andCouchman, J. R. (2015) Cell surface heparan sulfate proteoglycans control adhesion and invasion of breast carcinoma cells Molecular Cancer **14**, 15
254. Lim, H. C., andCouchman, J. R. (2014) Syndecan-2 regulation of morphology in breast carcinoma cells is dependent on RhoGTPases Biochim Biophys Acta **1840**, 2482-2490
255. Loftus, P. G., Watson, L., Deedigan, L. M., Camarillo-Retamosa, E., Dwyer, R. M., O'Flynn, L. *et al.* (2021) Targeting stromal cell Syndecan-2 reduces breast tumour growth, metastasis and limits immune evasion Int J Cancer **148**, 1245-1259
256. Yao, J., Li, W. Y., Li, S. G., Feng, X. S., andGao, S. G. (2014) Midkine promotes perineural invasion in human pancreatic cancer World J Gastroenterol **20**, 3018-3024
257. Hillemeyer, L., Espinoza-Sanchez, N. A., Greve, B., Hassan, N., Chelariu-Raicu, A., Kiesel, L. *et al.* (2022) The Cell Surface Heparan Sulfate Proteoglycan Syndecan-3 Promotes Ovarian Cancer Pathogenesis Int J Mol Sci **23**,
258. Prieto-Fernández, E., Egia-Mendikute, L., Bosch, A., García Del Río, A., Jimenez-Lasheras, B., Antoñana-Vildosola, A. *et al.* (2020) Hypoxia Promotes Syndecan-3 Expression in the Tumor Microenvironment Front Immunol **11**, 586977
259. Onyeisi, J. O. S., Lopes, C. C., andGötte, M. (2021) Syndecan-4 as a Pathogenesis Factor and Therapeutic Target in Cancer Biomolecules **11**, 503
260. Toba-Ichihashi, Y., Yamaoka, T., Ohmori, T., andOhba, M. (2016) Up-regulation of Syndecan-4 contributes to TGF-β1-induced epithelial to mesenchymal transition in lung adenocarcinoma A549 cells Biochem Biophys Rep **5**, 1-7
261. Na, K. Y., Bacchini, P., Bertoni, F., Kim, Y. W., andPark, Y. K. (2012) Syndecan-4 and fibronectin in osteosarcoma Pathology **44**, 325-330
262. Ma, M., Zhuang, J., Li, H., Mi, R., Song, Y., Yang, W. *et al.* (2024) Low expression of ZFP36L1 in osteosarcoma promotes lung metastasis by inhibiting the SDC4-TGF-β signaling feedback loop Oncogene **43**, 47-60
263. Charni, F., Friand, V., Haddad, O., Hlawaty, H., Martin, L., Vassy, R. *et al.* (2009) Syndecan-1 and syndecan-4 are involved in RANTES/CCL5-induced migration and invasion of human hepatoma cells Biochim Biophys Acta **1790**, 1314-1326
264. Labropoulou, V. T., Skandalis, S. S., Ravazoula, P., Perimenis, P., Karamanos, N. K., Kalofonos, H. P. *et al.* (2013) Expression of syndecan-4 and correlation with metastatic potential in testicular germ cell tumours Biomed Res Int **2013**, 214864
265. Jechorek, D., Haeusler-Pliske, I., Meyer, F., andRoessner, A. (2021) Diagnostic value of syndecan-4 protein expression in colorectal cancer Pathol Res Pract **222**, 153431
266. Takashima, S., Oka, Y., Fujiki, F., Morimoto, S., Nakajima, H., Nakae, Y. *et al.* (2016) Syndecan-4 as a biomarker to predict clinical outcome for glioblastoma multiforme treated with WT1 peptide vaccine Future Sci OA **2**, Fso96
267. Baba, F., Swartz, K., van Buren, R., Eickhoff, J., Zhang, Y., Wolberg, W. *et al.* (2006) Syndecan-1 and syndecan-4 are overexpressed in an estrogen receptor-negative, highly proliferative breast carcinoma subtype Breast Cancer Res Treat **98**, 91-98
268. Lendorf, M. E., Manon-Jensen, T., Kronqvist, P., Multhaupt, H. A. B., andCouchman, J. R. (2011) Syndecan-1 and Syndecan-4 Are Independent Indicators in Breast Carcinoma Journal of Histochemistry & Cytochemistry **59**, 615-629
269. Onyeisi, J. O. S., Greve, B., Espinoza-Sánchez, N. A., Kiesel, L., Lopes, C. C., andGötte, M. (2021) microRNA-140-3p modulates invasiveness, motility, and extracellular matrix adhesion of breast cancer cells by targeting syndecan-4 J Cell Biochem **122**, 1491-1505

270. Poças, J., Marques, C., Gomes, C., Otake, A. H., Pinto, F., Ferreira, M. *et al.* (2023) Syndecan-4 is a maestro of gastric cancer cell invasion and communication that underscores poor survival Proc Natl Acad Sci U S A **120**, e2214853120
271. Kleeff, J., Ishiwata, T., Kumbasar, A., Friess, H., Büchler, M. W., Lander, A. D. *et al.* (1998) The cell-surface heparan sulfate proteoglycan glypican-1 regulates growth factor action in pancreatic carcinoma cells and is overexpressed in human pancreatic cancer J Clin Invest **102**, 1662-1673
272. Duan, L., Hu, X. Q., Feng, D. Y., Lei, S. Y., andHu, G. H. (2013) GPC-1 may serve as a predictor of perineural invasion and a prognosticator of survival in pancreatic cancer Asian J Surg **36**, 7-12
273. Aikawa, T., Whipple, C. A., Lopez, M. E., Gunn, J., Young, A., Lander, A. D. *et al.* (2008) Glypican-1 modulates the angiogenic and metastatic potential of human and mouse cancer cells J Clin Invest **118**, 89-99
274. Melo, S. A., Luecke, L. B., Kahlert, C., Fernandez, A. F., Gammon, S. T., Kaye, J. *et al.* (2015) Glypican-1 identifies cancer exosomes and detects early pancreatic cancer Nature **523**, 177-182
275. Matsuda, K., Maruyama, H., Guo, F., Kleeff, J., Itakura, J., Matsumoto, Y. *et al.* (2001) Glypican-1 is overexpressed in human breast cancer and modulates the mitogenic effects of multiple heparin-binding growth factors in breast cancer cells Cancer Res **61**, 5562-5569
276. Saito, T., Sugiyama, K., Hama, S., Yamasaki, F., Takayasu, T., Nosaka, R. *et al.* (2017) High Expression of Glypican-1 Predicts Dissemination and Poor Prognosis in Glioblastomas World Neurosurg **105**, 282-288
277. Li, N., Fu, H., Hewitt, S. M., Dimitrov, D. S., andHo, M. (2017) Therapeutically targeting glypican-2 via single-domain antibody-based chimeric antigen receptors and immunotoxins in neuroblastoma Proc Natl Acad Sci U S A **114**, E6623-e6631
278. Baumhoer, D., Tornillo, L., Stadlmann, S., Roncalli, M., Diamantis, E. K., andTerracciano, L. M. (2008) Glypican 3 expression in human nonneoplastic, preneoplastic, and neoplastic tissues: a tissue microarray analysis of 4,387 tissue samples Am J Clin Pathol **129**, 899-906
279. Capurro, M. I., Xiang, Y. Y., Lobe, C., andFilmus, J. (2005) Glypican-3 promotes the growth of hepatocellular carcinoma by stimulating canonical Wnt signaling Cancer Res **65**, 6245-6254
280. Cao, J., Ma, J., Sun, L., Li, J., Qin, T., Zhou, C. *et al.* (2018) Targeting glypican-4 overcomes 5-FU resistance and attenuates stem cell-like properties via suppression of Wnt/β-catenin pathway in pancreatic cancer cells J Cell Biochem **119**, 9498-9512
281. Li, F., Shi, W., Capurro, M., andFilmus, J. (2011) Glypican-5 stimulates rhabdomyosarcoma cell proliferation by activating Hedgehog signaling J Cell Biol **192**, 691-704
282. Karapetsas, A., Giannakakis, A., Dangaj, D., Lanitis, E., Kynigopoulos, S., Lambropoulou, M. *et al.* (2015) Overexpression of GPC6 and TMEM132D in Early Stage Ovarian Cancer Correlates with CD8+ T-Lymphocyte Infiltration and Increased Patient Survival Biomed Res Int **2015**, 712438
283. Zhang, C., Liu, Z., Wang, L., Qiao, B., Du, E., Li, L. *et al.* (2016) Prognostic significance of GPC5 expression in patients with prostate cancer Tumour Biol **37**, 6413-6418
284. Karibe, T., Fukui, H., Sekikawa, A., Shiratori, K., andFujimori, T. (2008) EXTL3 promoter methylation down-regulates EXTL3 and heparan sulphate expression in mucinous colorectal cancers J Pathol **216**, 32-42
285. Sembajwe, L. F., Katta, K., Grønning, M., andKusche-Gullberg, M. (2018) The exostosin family of glycosyltransferases: mRNA expression profiles and heparan sulphate structure in human breast carcinoma cell lines Biosci Rep **38**,
286. Fernández-Vega, I., García-Suárez, O., García, B., Crespo, A., Astudillo, A., andQuirós, L. M. (2015) Heparan sulfate proteoglycans undergo differential expression alterations in right sided colorectal cancer, depending on their metastatic character BMC Cancer **15**, 742
287. Crespo, A., García-Suárez, O., Fernández-Vega, I., Solis-Hernandez, M. P., García, B., Castañón, S. *et al.* (2018) Heparan sulfate proteoglycans undergo differential expression

alterations in left sided colorectal cancer, depending on their metastatic character BMC Cancer **18**, 687
288. Tzeng, S. T., Tsai, M. H., Chen, C. L., Lee, J. X., Jao, T. M., Yu, S. L. *et al.* (2013) NDST4 is a novel candidate tumor suppressor gene at chromosome 4q26 and its genetic loss predicts adverse prognosis in colorectal cancer PLoS One **8**, e67040
289. Fuster, M. M., Wang, L., Castagnola, J., Sikora, L., Reddi, K., Lee, P. H. *et al.* (2007) Genetic alteration of endothelial heparan sulfate selectively inhibits tumor angiogenesis J Cell Biol **177**, 539-549
290. Grigorieva, E., Eshchenko, T., Rykova, V. I., Chernakov, A., Zabarovsky, E., andSidorov, S. V. (2008) Decreased expression of human D-glucuronyl C5-epimerase in breast cancer Int J Cancer **122**, 1172-1176
291. Grigorieva, E. V., Prudnikova, T. Y., Domanitskaya, N. V., Mostovich, L. A., Pavlova, T. V., Kashuba, V. I. *et al.* (2011) D-glucuronyl C5-epimerase suppresses small-cell lung cancer cell proliferation in vitro and tumour growth in vivo Br J Cancer **105**, 74-82
292. Prudnikova, T. Y., Mostovich, L. A., Domanitskaya, N. V., Pavlova, T. V., Kashuba, V. I., Zabarovsky, E. R. *et al.* (2010) Antiproliferative effect of D-glucuronyl C5-epimerase in human breast cancer cells Cancer Cell Int **10**, 27
293. Ferguson, B. W., andDatta, S. (2011) Role of heparan sulfate 2-o-sulfotransferase in prostate cancer cell proliferation, invasion, and growth factor signaling Prostate Cancer **2011**, 893208
294. Zhao, S., andWang, Z. (2020) Changes in heparan sulfate sulfotransferases and cell-surface heparan sulfate during SKM-1 cells granulocytic differentiation and A549 cells epithelial-mesenchymal transition Glycoconj J **37**, 151-164
295. Teixeira, F., Vijaya Kumar, A., Kumar Katakam, S., Cocola, C., Pelucchi, P., Graf, M. *et al.* (2020) The Heparan Sulfate Sulfotransferases HS2ST1 and HS3ST2 Are Novel Regulators of Breast Cancer Stem-Cell Properties Front Cell Dev Biol **8**, 559554
296. Vijaya Kumar, A., Brézillon, S., Untereiner, V., Sockalingum, G. D., Kumar Katakam, S., Mohamed, H. T. *et al.* (2020) HS2ST1-dependent signaling pathways determine breast cancer cell viability, matrix interactions, and invasive behavior Cancer Sci **111**, 2907-2922
297. Waaijer, C. J., de Andrea, C. E., Hamilton, A., van Oosterwijk, J. G., Stringer, S. E., andBovée, J. V. (2012) Cartilage tumour progression is characterized by an increased expression of heparan sulphate 6O-sulphation-modifying enzymes Virchows Arch **461**, 475-481
298. Hatabe, S., Kimura, H., Arao, T., Kato, H., Hayashi, H., Nagai, T. *et al.* (2013) Overexpression of heparan sulfate 6-O-sulfotransferase-2 in colorectal cancer Mol Clin Oncol **1**, 845-850
299. Iravani, O., Bay, B. H., andYip, G. W. (2017) Silencing HS6ST3 inhibits growth and progression of breast cancer cells through suppressing IGF1R and inducing XAF1 Exp Cell Res **350**, 380-389
300. Ferreras, C., Rushton, G., Cole, C. L., Babur, M., Telfer, B. A., van Kuppevelt, T. H. *et al.* (2012) Endothelial heparan sulfate 6-O-sulfation levels regulate angiogenic responses of endothelial cells to fibroblast growth factor 2 and vascular endothelial growth factor J Biol Chem **287**, 36132-36146
301. Cole, C. L., Rushton, G., Jayson, G. C., andAvizienyte, E. (2014) Ovarian cancer cell heparan sulfate 6-O-sulfotransferases regulate an angiogenic program induced by heparin-binding epidermal growth factor (EGF)-like growth factor/EGF receptor signaling J Biol Chem **289**, 10488-10501
302. Miyamoto, K., Asada, K., Fukutomi, T., Okochi, E., Yagi, Y., Hasegawa, T. *et al.* (2003) Methylation-associated silencing of heparan sulfate D-glucosaminyl 3-O-sulfotransferase-2 (3-OST-2) in human breast, colon, lung and pancreatic cancers Oncogene **22**, 274-280
303. Vijaya Kumar, A., Salem Gassar, E., Spillmann, D., Stock, C., Sen, Y. P., Zhang, T. *et al.* (2014) HS3ST2 modulates breast cancer cell invasiveness via MAP kinase- and Tcf4 (Tcf7l2)-dependent regulation of protease and cadherin expression Int J Cancer **135**, 2579-2592

304. Hwang, J. A., Kim, Y., Hong, S. H., Lee, J., Cho, Y. G., Han, J. Y. *et al.* (2013) Epigenetic inactivation of heparan sulfate (glucosamine) 3-O-sulfotransferase 2 in lung cancer and its role in tumorigenesis PLoS One **8**, e79634
305. Mao, X., Gauche, C., Coughtrie, M. W., Bui, C., Gulberti, S., Merhi-Soussi, F. *et al.* (2016) The heparan sulfate sulfotransferase 3-OST3A (HS3ST3A) is a novel tumor regulator and a prognostic marker in breast cancer Oncogene **35**, 5043-5055
306. Zhang, L., Song, K., Zhou, L., Xie, Z., Zhou, P., Zhao, Y. *et al.* (2015) Heparan sulfate D-glucosaminyl 3-O-sulfotransferase-3B1 (HS3ST3B1) promotes angiogenesis and proliferation by induction of VEGF in acute myeloid leukemia cells J Cell Biochem **116**, 1101-1112
307. Zhang, Z., Jiang, H., Wang, Y., andShi, M. (2018) Heparan sulfate D-glucosamine 3-O-sulfotransferase 3B1 is a novel regulator of transforming growth factor-beta-mediated epithelial-to-mesenchymal transition and regulated by miR-218 in nonsmall cell lung cancer J Cancer Res Ther **14**, 24-29
308. Cherfils-Vicini, J., Iltis, C., Cervera, L., Pisano, S., Croce, O., Sadouni, N. *et al.* (2019) Cancer cells induce immune escape via glycocalyx changes controlled by the telomeric protein TRF2 Embo j **38**,
309. Biroccio, A., Cherfils-Vicini, J., Augereau, A., Pinte, S., Bauwens, S., Ye, J. *et al.* (2013) TRF2 inhibits a cell-extrinsic pathway through which natural killer cells eliminate cancer cells Nat Cell Biol **15**, 818-828
310. Tátrai, P., Egedi, K., Somorácz, A., van Kuppevelt, T. H., Ten Dam, G., Lyon, M. *et al.* (2010) Quantitative and qualitative alterations of heparan sulfate in fibrogenic liver diseases and hepatocellular cancer J Histochem Cytochem **58**, 429-441
311. Ucakturk, E., Akman, O., Sun, X., Baydar, D. E., Dolgun, A., Zhang, F. *et al.* (2016) Changes in composition and sulfation patterns of glycoaminoglycans in renal cell carcinoma Glycoconj J **33**, 103-112
312. Hammond, E., Khurana, A., Shridhar, V., andDredge, K. (2014) The Role of Heparanase and Sulfatases in the Modification of Heparan Sulfate Proteoglycans within the Tumor Microenvironment and Opportunities for Novel Cancer Therapeutics Front Oncol **4**, 195
313. Vivès, R. R., Seffouh, A., andLortat-Jacob, H. (2014) Post-Synthetic Regulation of HS Structure: The Yin and Yang of the Sulfs in Cancer Front Oncol **3**, 331
314. Lai, J. P., Chien, J., Strome, S. E., Staub, J., Montoya, D. P., Greene, E. L. *et al.* (2004) HSulf-1 modulates HGF-mediated tumor cell invasion and signaling in head and neck squamous carcinoma Oncogene **23**, 1439-1447
315. Lai, J. P., Chien, J. R., Moser, D. R., Staub, J. K., Aderca, I., Montoya, D. P. *et al.* (2004) hSulf1 Sulfatase promotes apoptosis of hepatocellular cancer cells by decreasing heparin-binding growth factor signaling Gastroenterology **126**, 231-248
316. Li, J., Kleeff, J., Abiatari, I., Kayed, H., Giese, N. A., Felix, K. *et al.* (2005) Enhanced levels of Hsulf-1 interfere with heparin-binding growth factor signaling in pancreatic cancer Mol Cancer **4**, 14
317. Bret, C., Moreaux, J., Schved, J. F., Hose, D., andKlein, B. (2011) SULFs in human neoplasia: implication as progression and prognosis factors J Transl Med **9**, 72
318. Lemjabbar-Alaoui, H., van Zante, A., Singer, M. S., Xue, Q., Wang, Y. Q., Tsay, D. *et al.* (2010) Sulf-2, a heparan sulfate endosulfatase, promotes human lung carcinogenesis Oncogene **29**, 635-646
319. Nawroth, R., van Zante, A., Cervantes, S., McManus, M., Hebrok, M., andRosen, S. D. (2007) Extracellular sulfatases, elements of the Wnt signaling pathway, positively regulate growth and tumorigenicity of human pancreatic cancer cells PLoS One **2**, e392
320. Morimoto-Tomita, M., Uchimura, K., Bistrup, A., Lum, D. H., Egeblad, M., Boudreau, N. *et al.* (2005) Sulf-2, a Proangiogenic Heparan Sulfate Endosulfatase, Is Upregulated in Breast Cancer Neoplasia **7**, 1001-1010

321. Tao, Y., Han, T., Zhang, T., andSun, C. (2017) Sulfatase-2 promotes the growth and metastasis of colorectal cancer by activating Akt and Erk1/2 pathways Biomedicine & Pharmacotherapy **89**, 1370-1377
322. Lai, J. P., Sandhu, D. S., Yu, C., Han, T., Moser, C. D., Jackson, K. K. *et al.* (2008) Sulfatase 2 up-regulates glypican 3, promotes fibroblast growth factor signaling, and decreases survival in hepatocellular carcinoma Hepatology **47**, 1211-1222
323. El Masri, R., Seffouh, A., Roelants, C., Seffouh, I., Gout, E., Pérard, J. *et al.* (2022) Extracellular endosulfatase Sulf-2 harbors a chondroitin/dermatan sulfate chain that modulates its enzyme activity Cell Reports **38**, 110516
324. Lo, S. L., Thike, A. A., Tan, S. Y., Lim, T. K., Tan, I. B., Choo, S. P. *et al.* (2011) Expression of heparan sulfate in gastric carcinoma and its correlation with clinicopathological features and patient survival J Clin Pathol **64**, 153-158
325. Weyers, A., Yang, B., Park, J. H., Kim, Y. S., Kim, S. M., Lee, S. E. *et al.* (2013) Microanalysis of stomach cancer glycosaminoglycans Glycoconj J **30**, 701-707
326. Wiksten, J. P., Lundin, J., Nordling, S., Lundin, M., Kokkola, A., von Boguslawski, K. *et al.* (2001) Epithelial and stromal syndecan-1 expression as predictor of outcome in patients with gastric cancer Int J Cancer **95**, 1-6
327. Magalhães, A., Marcos, N. T., Carvalho, A. S., David, L., Figueiredo, C., Bastos, J. *et al.* (2009) Helicobacter pylori cag pathogenicity island-positive strains induce syndecan-4 expression in gastric epithelial cells FEMS Immunol Med Microbiol **56**, 223-232
328. Dinccelik-Aslan, M., Gumus-Akay, G., Elhan, A. H., Unal, E., andTukun, A. (2015) Diagnostic and prognostic significance of glypican 5 and glypican 6 gene expression levels in gastric adenocarcinoma Mol Clin Oncol **3**, 584-590
329. Pinto, F., Santos-Ferreira, L., Pinto, M. T., Gomes, C., andReis, C. A. (2021) The Extracellular Small Leucine-Rich Proteoglycan Biglycan Is a Key Player in Gastric Cancer Aggressiveness Cancers (Basel) **13**,
330. Shi, X., Su, S., Long, J., Mei, B., andChen, Y. (2011) MicroRNA-191 targets N-deacetylase/N-sulfotransferase 1 and promotes cell growth in human gastric carcinoma cell line MGC803 Acta Biochim Biophys Sin (Shanghai) **43**, 849-856
331. Jin, Y., He, J., Du, J., Zhang, R. X., Yao, H. B., andShao, Q. S. (2017) Overexpression of HS6ST2 is associated with poor prognosis in patients with gastric cancer Oncol Lett **14**, 6191-6197
332. Bernal, C., Aguayo, F., Villarroel, C., Vargas, M., Díaz, I., Ossandon, F. J. *et al.* (2008) Reprimo as a potential biomarker for early detection in gastric cancer Clin Cancer Res **14**, 6264-6269
333. Chen, Z., Fan, J. Q., Li, J., Li, Q. S., Yan, Z., Jia, X. K. *et al.* (2009) Promoter hypermethylation correlates with the Hsulf-1 silencing in human breast and gastric cancer Int J Cancer **124**, 739-744
334. Li, J., Mo, M. L., Chen, Z., Yang, J., Li, Q. S., Wang, D. J. *et al.* (2011) HSulf-1 inhibits cell proliferation and invasion in human gastric cancer Cancer Sci **102**, 1815-1821
335. Hur, K., Han, T.-S., Jung, E.-J., Yu, J., Lee, H.-J., Kim, W. H. *et al.* (2012) Up-regulated expression of sulfatases (SULF1 and SULF2) as prognostic and metastasis predictive markers in human gastric cancer The Journal of Pathology **228**, 88-98
336. Takaoka, M., Naomoto, Y., Ohkawa, T., Uetsuka, H., Shirakawa, Y., Uno, F. *et al.* (2003) Heparanase Expression Correlates with Invasion and Poor Prognosis in Gastric Cancers Laboratory Investigation **83**, 613-622
337. Byrne, D. P., Li, Y., Ramakrishnan, K., Barsukov, I. L., Yates, E. A., Eyers, C. E. *et al.* (2018) New tools for carbohydrate sulfation analysis: heparan sulfate 2-O-sulfotransferase (HS2ST) is a target for small-molecule protein kinase inhibitors Biochem J **475**, 2417-2433
338. Maciej-Hulme, M. L., Dubaissi, E., Shao, C., Zaia, J., Amaya, E., Flitsch, S. L. *et al.* (2021) Selective Inhibition of Heparan Sulphate and Not Chondroitin Sulphate Biosynthesis by a Small, Soluble Competitive Inhibitor Int J Mol Sci **22**,

339. Raman, K., andKuberan, B. (2010) Click-xylosides mitigate glioma cell invasion in vitro Mol Biosyst **6**, 1800-1802
340. Raman, K., Ninomiya, M., Nguyen, T. K., Tsuzuki, Y., Koketsu, M., andKuberan, B. (2011) Novel glycosaminoglycan biosynthetic inhibitors affect tumor-associated angiogenesis Biochem Biophys Res Commun **404**, 86-89
341. Kalita, M., Villanueva-Meyer, J., Ohkawa, Y., Kalyanaraman, C., Chen, K., Mohamed, E. *et al.* (2021) Synthesis and Screening of α-Xylosides in Human Glioblastoma Cells Mol Pharm **18**, 451-460
342. Morla, S. (2019) Glycosaminoglycans and Glycosaminoglycan Mimetics in Cancer and Inflammation Int J Mol Sci **20**,
343. Gao, W., Kim, H., andHo, M. (2015) Human Monoclonal Antibody Targeting the Heparan Sulfate Chains of Glypican-3 Inhibits HGF-Mediated Migration and Motility of Hepatocellular Carcinoma Cells PLoS One **10**, e0137664
344. Yu, T., Chaganty, B., Lin, L., Xing, L., Ramakrishnan, B., Wen, K. *et al.* (2020) VIS832, a novel CD138-targeting monoclonal antibody, potently induces killing of human multiple myeloma and further synergizes with IMiDs or bortezomib in vitro and in vivo Blood Cancer J **10**, 110
345. Afratis, N., Gialeli, C., Nikitovic, D., Tsegenidis, T., Karousou, E., Theocharis, A. D. *et al.* (2012) Glycosaminoglycans: key players in cancer cell biology and treatment The FEBS journal **279**, 1177-1197
346. Xu, D., andEsko, J. D. (2014) Demystifying heparan sulfate-protein interactions Annual review of biochemistry **83**, 129-157
347. Nikitovic, D., Berdiaki, A., Spyridaki, I., Krasanakis, T., Tsatsakis, A., andTzanakakis, G. N. (2018) Proteoglycans-Biomarkers and Targets in Cancer Therapy Front Endocrinol (Lausanne) **9**, 69
348. Pönighaus, C., Ambrosius, M., Casanova, J. C., Prante, C., Kuhn, J., Esko, J. D. *et al.* (2007) Human xylosyltransferase II is involved in the biosynthesis of the uniform tetrasaccharide linkage region in chondroitin sulfate and heparan sulfate proteoglycans J Biol Chem **282**, 5201-5206
349. Marques, C., Reis, C. A., Vivès, R. R., andMagalhães, A. (2021) Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression Frontiers in oncology **11**, 778752
350. Kim, B. T., Kitagawa, H., Tanaka, J., Tamura, J., andSugahara, K. (2003) In vitro heparan sulfate polymerization: crucial roles of core protein moieties of primer substrates in addition to the EXT1-EXT2 interaction J Biol Chem **278**, 41618-41623
351. Grobe, K., Ledin, J., Ringvall, M., Holmborn, K., Forsberg, E., Esko, J. D. *et al.* (2002) Heparan sulfate and development: differential roles of the N-acetylglucosamine N-deacetylase/N-sulfotransferase isozymes Biochimica et Biophysica Acta (BBA) - General Subjects **1573**, 209-215
352. Busse-Wicher, M., Wicher, K. B., andKusche-Gullberg, M. (2014) The extostosin family: Proteins with many functions Matrix Biology **35**, 25-33
353. Wilson, L. F. L., Dendooven, T., Hardwick, S. W., Echevarría-Poza, A., Tryfona, T., Krogh, K. *et al.* (2022) The structure of EXTL3 helps to explain the different roles of bi-domain exostosins in heparan sulfate synthesis Nat Commun **13**, 3314
354. Takahashi, I., Noguchi, N., Nata, K., Yamada, S., Kaneiwa, T., Mizumoto, S. *et al.* (2009) Important role of heparan sulfate in postnatal islet growth and insulin secretion Biochemical and biophysical research communications **383**, 113-118
355. Lee, J. S., von der Hardt, S., Rusch, M. A., Stringer, S. E., Stickney, H. L., Talbot, W. S. *et al.* (2004) Axon sorting in the optic tract requires HSPG synthesis by ext2 (dackel) and extl3 (boxer) Neuron **44**, 947-960
356. Chen, Y. H., Narimatsu, Y., Clausen, T. M., Gomes, C., Karlsson, R., Steentoft, C. *et al.* (2018) The GAGOme: a cell-based library of displayed glycosaminoglycans Nat Methods **15**, 881-888

357. Busse, M., Feta, A., Presto, J., Wilen, M., Gronning, M., Kjellen, L. *et al.* (2007) Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation J Biol Chem **282**, 32802-32810
358. Purnomo, E., Emoto, N., Nugrahaningsih, D. A., Nakayama, K., Yagi, K., Heiden, S. *et al.* (2013) Glycosaminoglycan overproduction in the aorta increases aortic calcification in murine chronic kidney disease Journal of the American Heart Association **2**, e000405
359. Katta, K., Imran, T., Busse-Wicher, M., Grønning, M., Czajkowski, S., andKusche-Gullberg, M. (2015) Reduced Expression of EXTL2, a Member of the Exostosin (EXT) Family of Glycosyltransferases, in Human Embryonic Kidney 293 Cells Results in Longer Heparan Sulfate Chains J Biol Chem **290**, 13168-13177
360. Okada, M., Nadanaka, S., Shoji, N., Tamura, J., andKitagawa, H. (2010) Biosynthesis of heparan sulfate in EXT1-deficient cells The Biochemical journal **428**, 463-471
361. Gingis-Velitski, S., Zetser, A., Kaplan, V., Ben-Zaken, O., Cohen, E., Levy-Adam, F. *et al.* (2004) Heparanase uptake is mediated by cell membrane heparan sulfate proteoglycans J Biol Chem **279**, 44084-44092
362. Gopal, S., Bober, A., Whiteford, J. R., Multhaupt, H. A., Yoneda, A., andCouchman, J. R. (2010) Heparan sulfate chain valency controls syndecan-4 function in cell adhesion J Biol Chem **285**, 14247-14258
363. Lidholt, K., Weinke, J. L., Kiser, C. S., Lugemwa, F. N., Bame, K. J., Cheifetz, S. *et al.* (1992) A single mutation affects both N-acetylglucosaminyltransferase and glucuronosyltransferase activities in a Chinese hamster ovary cell mutant defective in heparan sulfate biosynthesis Proceedings of the National Academy of Sciences of the United States of America **89**, 2267-2271
364. Holmborn, K., Habicher, J., Kasza, Z., Eriksson, A. S., Filipek-Gorniok, B., Gopal, S. *et al.* (2012) On the roles and regulation of chondroitin sulfate and heparan sulfate in zebrafish pharyngeal cartilage morphogenesis J Biol Chem **287**, 33905-33916
365. Izumikawa, T., Sato, B., Mikami, T., Tamura, J., Igarashi, M., andKitagawa, H. (2015) GlcUAβ1-3Galβ1-3Galβ1-4Xyl(2-O-phosphate) is the preferred substrate for chondroitin N-acetylgalactosaminyltransferase-1 J Biol Chem **290**, 5438-5448
366. Depau, L., Brunetti, J., Falciani, C., Mandarini, E., Riolo, G., Zanchi, M. *et al.* (2020) Heparan Sulfate Proteoglycans Can Promote Opposite Effects on Adhesion and Directional Migration of Different Cancer Cells Journal of medicinal chemistry **63**, 15997-16011
367. Li, N., Spetz, M. R., andHo, M. (2020) The Role of Glypicans in Cancer Progression and Therapy Journal of Histochemistry & Cytochemistry **68**, 841-862
368. Li, N., Gao, W., Zhang, Y. F., andHo, M. (2018) Glypicans as Cancer Therapeutic Targets Trends Cancer **4**, 741-754
369. Le Jan, S., Hayashi, M., Kasza, Z., Eriksson, I., Bishop, J. R., Weibrecht, I. *et al.* (2012) Functional overlap between chondroitin and heparan sulfate proteoglycans during VEGF-induced sprouting angiogenesis Arteriosclerosis, thrombosis, and vascular biology **32**, 1255-1263
370. Shintani, Y., Takashima, S., Asano, Y., Kato, H., Liao, Y., Yamazaki, S. *et al.* (2006) Glycosaminoglycan modification of neuropilin-1 modulates VEGFR2 signaling The EMBO journal **25**, 3045-3055
371. Lanzi, C., andCassinelli, G. (2020) Receptor tyrosine kinases and heparan sulfate proteoglycans: Interplay providing anticancer targeting strategies and new therapeutic opportunities Biochemical pharmacology **178**, 114084
372. Kania, A., andKlein, R. (2016) Mechanisms of ephrin-Eph signalling in development, physiology and disease Nature reviews Molecular cell biology **17**, 240-256
373. Fukai, J., Yokote, H., Yamanaka, R., Arao, T., Nishio, K., andItakura, T. (2008) EphA4 promotes cell proliferation and migration through a novel EphA4-FGFR1 signaling pathway in the human glioma U251 cell line Molecular cancer therapeutics **7**, 2768-2778

374. Furuhashi, S., Morita, Y., Ida, S., Muraki, R., Kitajima, R., Takeda, M. *et al.* (2021) Ephrin Receptor A4 Expression Enhances Migration, Invasion and Neurotropism in Pancreatic Ductal Adenocarcinoma Cells Anticancer Res **41**, 1733-1744
375. Oki, M., Yamamoto, H., Taniguchi, H., Adachi, Y., Imai, K., andShinomura, Y. (2008) Overexpression of the receptor tyrosine kinase EphA4 in human gastric cancers World journal of gastroenterology **14**, 5650-5656
376. Miyazaki, K., Inokuchi, M., Takagi, Y., Kato, K., Kojima, K., andSugihara, K. (2013) EphA4 is a prognostic factor in gastric cancer BMC Clinical Pathology **13**, 19
377. Inokuchi, M., Nakagawa, M., Baogok, N., Takagi, Y., Tanioka, T., Gokita, K. *et al.* (2018) Prognostic Significance of High EphA1-4 Expression Levels in Locally Advanced Gastric Cancer Anticancer Research **38**, 1685-1693
378. Irie, F., Okuno, M., Matsumoto, K., Pasquale, E. B., andYamaguchi, Y. (2008) Heparan sulfate regulates ephrin-A3/EphA receptor signaling Proceedings of the National Academy of Sciences of the United States of America **105**, 12307-12312
379. Corti, F., Wang, Y., Rhodes, J. M., Atri, D., Archer-Hartmann, S., Zhang, J. *et al.* (2019) N-terminal syndecan-2 domain selectively enhances 6-O heparan sulfate chains sulfation and promotes VEGFA165-dependent neovascularization Nat Commun **10**, 1562
380. Ohkawa, Y., Wade, A., Lindberg, O. R., Chen, K. Y., Tran, V. M., Brown, S. J. *et al.* (2021) Heparan Sulfate Synthesized by Ext1 Regulates Receptor Tyrosine Kinase Signaling and Promotes Resistance to EGFR Inhibitors in GBM Molecular Cancer Research **19**, 150-161
381. Arai, T., Akiyama, Y., Nagasaki, H., Murase, N., Okabe, S., Ikeuchi, T. *et al.* (1999) EXTL3/EXTR1 alterations in colorectal cancer cell lines International journal of oncology **15**, 915-919
382. Nadanaka, S., Hashiguchi, T., andKitagawa, H. (2020) Aberrant glycosaminoglycan biosynthesis by tumor suppressor EXTL2 deficiency promotes liver inflammation and tumorigenesis through Toll-like 4 receptor signaling FASEB journal : official publication of the Federation of American Societies for Experimental Biology **34**, 8385-8401
383. Lonowski, L. A., Narimatsu, Y., Riaz, A., Delay, C. E., Yang, Z., Niola, F. *et al.* (2017) Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis Nat Protoc **12**, 581-603
384. Brinkman, E. K., Chen, T., Amendola, M., andvan Steensel, B. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition Nucleic Acids Res **42**, e168
385. Narimatsu, Y., Joshi, H. J., Yang, Z., Gomes, C., Chen, Y. H., Lorenzetti, F. C. *et al.* (2018) A validated gRNA library for CRISPR/Cas9 targeting of the human glycosyltransferase genome Glycobiology **28**, 295-305
386. David, G., van der Schueren, B., Marynen, P., Cassiman, J. J., andvan den Berghe, H. (1992) Molecular cloning of amphiglycan, a novel integral membrane heparan sulfate proteoglycan expressed by epithelial and fibroblastic cells J Cell Biol **118**, 961-969
387. Henriet, E., Jäger, S., Tran, C., Bastien, P., Michelet, J. F., Minondo, A. M. *et al.* (2017) A jasmonic acid derivative improves skin healing and induces changes in proteoglycan expression and glycosaminoglycan structure Biochimica et biophysica acta General subjects **1861**, 2250-2260
388. Marques, C., Poças, J., Gomes, C., Faria-Ramos, I., Reis, C. A., Vivès, R. R. *et al.* (2022) Glycosyltransferases EXTL2 and EXTL3 cellular balance dictates heparan sulfate biosynthesis and shapes gastric cancer cell motility and invasion J Biol Chem **298**, 102546
389. Bass, M. D., Roach, K. A., Morgan, M. R., Mostafavi-Pour, Z., Schoen, T., Muramatsu, T. *et al.* (2007) Syndecan-4-dependent Rac1 regulation determines directional migration in response to the extracellular matrix J Cell Biol **177**, 527-538
390. Lamanna, W. C., Baldwin, R. J., Padva, M., Kalus, I., Ten Dam, G., van Kuppevelt, T. H. *et al.* (2006) Heparan sulfate 6-O-endosulfatases: discrete in vivo activities and functional co-operativity Biochem J **400**, 63-73

391. Deligny, A., Dierker, T., Dagälv, A., Lundequist, A., Eriksson, I., Nairn, A. V. *et al.* (2016) NDST2 (N-Deacetylase/N-Sulfotransferase-2) Enzyme Regulates Heparan Sulfate Chain Length J Biol Chem **291**, 18600-18607
392. Lidholt, K., andLindahl, U. (1992) Biosynthesis of heparin. The D-glucuronosyl- and N-acetyl-D-glucosaminyltransferase reactions and their relation to polymer modification Biochem J **287 (Pt 1)**, 21-29
393. Laguri, C., Sadir, R., Gout, E., Vivès, R. R., andLortat-Jacob, H. (2022) Preparation and Characterization of Heparan Sulfate-Derived Oligosaccharides to Investigate Protein-GAG Interaction and HS Biosynthesis Enzyme Activity Methods Mol Biol **2303**, 121-137
394. Seffouh, A., El Masri, R., Makshakova, O., Gout, E., Hassoun, Z. E. O., Andrieu, J. P. *et al.* (2019) Expression and purification of recombinant extracellular sulfatase HSulf-2 allows deciphering of enzyme sub-domain coordinated role for the binding and 6-O-desulfation of heparan sulfate Cell Mol Life Sci **76**, 1807-1819
395. Lin, X., Han, T., Xia, Q., Cui, J., Zhuo, M., Liang, Y. *et al.* (2021) CHPF promotes gastric cancer tumorigenesis through the activation of E2F1 Cell Death & Disease **12**, 876
396. Liu, J., Tian, Z., Liu, T., Wen, D., Ma, Z., Liu, Y. *et al.* (2021) CHSY1 is upregulated and acts as tumor promotor in gastric cancer through regulating cell proliferation, apoptosis, and migration Cell Cycle **20**, 1861-1874
397. Huang, X., Liu, Y., Qian, C., Shen, Q., Wu, M., Zhu, B. *et al.* (2023) CHSY3 promotes proliferation and migration in gastric cancer and is associated with immune infiltration J Transl Med **21**, 474
398. da Cunha, C. B., Oliveira, C., Wen, X., Gomes, B., Sousa, S., Suriano, G. *et al.* (2010) De novo expression of CD44 variants in sporadic and hereditary gastric cancer Lab Invest **90**, 1604-1614
399. Lourenço, B. N., Springer, N. L., Ferreira, D., Oliveira, C., Granja, P. L., andFischbach, C. (2018) CD44v6 increases gastric cancer malignant phenotype by modulating adipose stromal cell-mediated ECM remodeling Integr Biol (Camb) **10**, 145-158
400. Mereiter, S., Martins Á, M., Gomes, C., Balmaña, M., Macedo, J. A., Polom, K. *et al.* (2019) O-glycan truncation enhances cancer-related functions of CD44 in gastric cancer FEBS Lett **593**, 1675-1689
401. Wang, Z., Patel, V. N., Song, X., Xu, Y., Kaminski, A. M., Doan, V. U. *et al.* (2023) Increased 3-O-sulfated heparan sulfate in Alzheimer's disease brain is associated with genetic risk gene HS3ST1 Sci Adv **9**, eadf6232
402. Pudełko, A., Wisowski, G., Olczyk, K., andKoźma, E. M. (2019) The dual role of the glycosaminoglycan chondroitin-6-sulfate in the development, progression and metastasis of cancer Febs j **286**, 1815-1837
403. Pikas, D. S., Eriksson, I., andKjellén, L. (2000) Overexpression of different isoforms of glucosaminyl N-deacetylase/N-sulfotransferase results in distinct heparan sulfate N-sulfation patterns Biochemistry **39**, 4552-4558
404. Liu, J., andPedersen, L. C. (2007) Anticoagulant heparan sulfate: structural specificity and biosynthesis Appl Microbiol Biotechnol **74**, 263-272
405. Suhovskih, A. V., Domanitskaya, N. V., Tsidulko, A. Y., Prudnikova, T. Y., Kashuba, V. I., andGrigorieva, E. V. (2015) Tissue-specificity of heparan sulfate biosynthetic machinery in cancer Cell Adh Migr **9**, 452-459
406. Pál, D., Tóth, G., Sugár, S., Fügedi, K. D., Szabó, D., Kovalszky, I. *et al.* (2023) Compositional Analysis of Glycosaminoglycans in Different Lung Cancer Types—A Pilot Study International Journal of Molecular Sciences **24**, 7050
407. Suhovskih, A. V., Tsidulko, A. Y., Kutsenko, O. S., Kovner, A. V., Aidagulova, S. V., Ernberg, I. *et al.* (2014) Transcriptional Activity of Heparan Sulfate Biosynthetic Machinery is Specifically Impaired in Benign Prostate Hyperplasia and Prostate Cancer Front Oncol **4**, 79
408. Hijmans, R. S., Shrestha, P., Sarpong, K. A., Yazdani, S., El Masri, R., de Jong, W. H. A. *et al.* (2017) High sodium diet converts renal proteoglycans into pro-inflammatory mediators in rats PLoS One **12**, e0178940

409. Lordick, F., Allum, W., Carneiro, F., Mitry, E., Tabernero, J., Tan, P. *et al.* (2014) Unmet needs and challenges in gastric cancer: the way forward Cancer Treat Rev **40**, 692-700
410. Nadanaka, S., andKitagawa, H. (2014) EXTL2 controls liver regeneration and aortic calcification through xylose kinase-dependent regulation of glycosaminoglycan biosynthesis Matrix Biol **35**, 18-24
411. Wang, S., Chen, N., Chen, Y., Sun, L., Li, L., andLiu, H. (2018) Elevated GPC3 level promotes cell proliferation in liver cancer Oncol Lett **16**, 970-976
412. Vicente, C. M., Lima, M. A., Nader, H. B., andToma, L. (2015) SULF2 overexpression positively regulates tumorigenicity of human prostate cancer cells J Exp Clin Cancer Res **34**, 25
413. Chalkiadaki, G., Nikitovic, D., Berdiaki, A., Sifaki, M., Krasagakis, K., Katonis, P. *et al.* (2009) Fibroblast growth factor-2 modulates melanoma adhesion and migration through a syndecan-4-dependent mechanism Int J Biochem Cell Biol **41**, 1323-1331
414. Zhang, L., Peng, K. W., Wang, B., Yang, X. F., andZhang, Z. M. (2020) EDIL3 regulates gastric cancer cell migration, invasion and epithelial-mesenchymal transition via TGF-β1/XIST/miR-137 feedback loop Transl Cancer Res **9**, 6313-6330
415. Ponandai-Srinivasan, S., Saare, M., Boggavarapu, N. R., Frisendahl, C., Ehrström, S., Riethmüller, C. *et al.* (2020) Syndecan-1 modulates the invasive potential of endometrioma via TGF-β signalling in a subgroup of women with endometriosis Hum Reprod **35**, 2280-2293
416. Dobra, K., Nurminen, M., andHjerpe, A. (2003) Growth factors regulate the expression profile of their syndecan co-receptors and the differentiation of mesothelioma cells Anticancer Res **23**, 2435-2444
417. Merry, C. L., Bullock, S. L., Swan, D. C., Backen, A. C., Lyon, M., Beddington, R. S. *et al.* (2001) The molecular phenotype of heparan sulfate in the Hs2st-/- mutant mouse J Biol Chem **276**, 35429-35434
418. Presto, J., Thuveson, M., Carlsson, P., Busse, M., Wilén, M., Eriksson, I. *et al.* (2008) Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation Proc Natl Acad Sci U S A **105**, 4751-4756
419. Pinhal, M. A., Smith, B., Olson, S., Aikawa, J., Kimata, K., andEsko, J. D. (2001) Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo Proc Natl Acad Sci U S A **98**, 12984-12989
420. Qin, Y., Ke, J., Gu, X., Fang, J., Wang, W., Cong, Q. *et al.* (2015) Structural and functional study of D-glucuronyl C5-epimerase J Biol Chem **290**, 4620-4630
421. Jaakkola, P., Määttä, A., andJalkanen, M. (1998) The activation and composition of FiRE (an FGF-inducible response element) differ in a cell type- and growth factor-specific manner Oncogene **17**, 1279-1286
422. Schönherr, E., Järveläinen, H. T., Kinsella, M. G., Sandell, L. J., andWight, T. N. (1993) Platelet-derived growth factor and transforming growth factor-beta 1 differentially affect the synthesis of biglycan and decorin by monkey arterial smooth muscle cells Arterioscler Thromb **13**, 1026-1036
423. Anggara, K., Sršan, L., Jaroentomeechai, T., Wu, X., Rauschenbach, S., Narimatsu, Y. *et al.* (2023) Direct observation of glycans bonded to proteins and lipids at the single-molecule level Science **382**, 219-223
424. Kunzke, T., Balluff, B., Feuchtinger, A., Buck, A., Langer, R., Luber, B. *et al.* (2017) Native glycan fragments detected by MALDI-FT-ICR mass spectrometry imaging impact gastric cancer biology and patient outcome Oncotarget **8**, 68012-68025
425. Purushothaman, A., Bandari, S. K., Chandrashekar, D. S., Jones, R. J., Lee, H. C., Weber, D. M. *et al.* (2017) Chondroitin sulfate proteoglycan serglycin influences protein cargo loading and functions of tumor-derived exosomes Oncotarget **8**, 73723-73732
426. Ghossoub, R., Chéry, M., Audebert, S., Leblanc, R., Egea-Jimenez, A. L., Lembo, F. *et al.* (2020) Tetraspanin-6 negatively regulates exosome production Proc Natl Acad Sci U S A **117**, 5913-5922

427. Cerezo-Magaña, M., Bång-Rudenstam, A., andBelting, M. (2020) The pleiotropic role of proteoglycans in extracellular vesicle mediated communication in the tumor microenvironment Seminars in Cancer Biology **62**, 99-107

UNIVERSIDADES PARTICIPANTES **UNIVERSIDADE DO PORTO**

UNIVERSITÉ GRENOBLE-ALPES